

## THE GRISM LENS-AMPLIFIED SURVEY FROM SPACE (GLASS) X. SUB-KPC RESOLUTION GAS-PHASE METALLICITY MAPS AT COSMIC NOON BEHIND THE HUBBLE FRONTIER FIELDS CLUSTER MACS1149.6+2223

XIN WANG<sup>1</sup>, TUCKER A. JONES<sup>2,3,18</sup>, TOMMASO TREU<sup>1</sup>, TAKAHIRO MORISHITA<sup>1,4,5</sup>, LOUIS E. ABRAMSON<sup>1</sup>, GABRIEL B. BRAMMER<sup>6</sup>, KUANG-HAN HUANG<sup>3</sup>, MATTHEW A. MALKAN<sup>1</sup>, KASPER B. SCHMIDT<sup>7</sup>, ADRIANO FONTANA<sup>8</sup>, CLAUDIO GRILLO<sup>9,10</sup>, ALAINA L. HENRY<sup>6,11</sup>, WOUTER KARMAN<sup>12</sup>, PATRICK L. KELLY<sup>13</sup>, CHARLOTTE A. MASON<sup>1</sup>, AMATA MERCURIO<sup>14</sup>, PIERO ROSATI<sup>15</sup>, KEREN SHARON<sup>16</sup>, MICHELE TRENTI<sup>17</sup>, BENEDETTA VULCANI<sup>17</sup><sup>1</sup> Department of Physics and Astronomy, University of California, Los Angeles, CA, USA 90095-1547<sup>2</sup> Institute of Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA<sup>3</sup> University of California Davis, 1 Shields Avenue, Davis, CA 95616, USA<sup>4</sup> Astronomical Institute, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan<sup>5</sup> Institute for International Advanced Research and Education, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan<sup>6</sup> Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD, 21218, USA<sup>7</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany<sup>8</sup> INAF - Osservatorio Astronomico di Roma Via Frascati 33 - 00040 Monte Porzio Catone, Italy<sup>9</sup> Dipartimento di Fisica, Università degli Studi di Milano, via Celoria 16, I-20133 Milano, Italy<sup>10</sup> Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, 2100 Copenhagen, Denmark<sup>11</sup> Astrophysics Science Division, Goddard Space Flight Center, Code 665, Greenbelt, MD 20771<sup>12</sup> Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands<sup>13</sup> Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA<sup>14</sup> INAF - Osservatorio Astronomico di Capodimonte, Via Moiraniello 16, I-80131 Napoli, Italy<sup>15</sup> Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, via Saragat 1, 44122 Ferrara, Italy<sup>16</sup> Department of Astronomy, University of Michigan, 1085 S. University Avenue, Ann Arbor, MI 48109, USA<sup>17</sup> School of Physics, University of Melbourne, VIC 3010, Australia and<sup>18</sup> Hubble Fellow

submitted to ApJ

## Abstract

We combine deep Hubble Space Telescope grism spectroscopy with a new Bayesian method to derive maps of gas-phase metallicity, nebular dust extinction, and star-formation rate for 10 star-forming galaxies at high redshift ( $1.2 \lesssim z \lesssim 2.3$ ). Exploiting lensing magnification by the foreground cluster MACS1149.6+2223, we reach sub-kpc spatial resolution and push the limit of stellar mass associated with such high- $z$  spatially resolved measurements below  $10^8 M_\odot$  for the first time. Our maps exhibit diverse morphologies, indicative of various effects such as efficient radial mixing from tidal torques, rapid accretion of low-metallicity gas, and other physical processes which can affect the gas and metallicity distributions in individual galaxies. Based upon an exhaustive sample of all existing sub-kpc resolution metallicity gradient measurements at high- $z$ , we find that predictions given by analytical chemical evolution models assuming a relatively extended star-formation profile in the early disk formation phase can explain the majority of observed metallicity gradients, without involving galactic feedback or radial outflows. We observe a tentative correlation between stellar mass and metallicity gradient, consistent with the “downsizing” galaxy formation picture that more massive galaxies are more evolved into a later phase of disk growth, where they experience more coherent mass assembly at all radii and thus show shallower metallicity gradients. In addition to the spatially resolved analysis, we compile a sample of homogeneously cross-calibrated integrated metallicity measurements spanning three orders of magnitude in stellar mass at  $z \sim 1.8$ . We use this sample to study the mass-metallicity relation and test whether star-formation rates drive the scatter in the mass-metallicity relation (i.e. the “fundamental metallicity relation”). The slope of the observed mass-metallicity relation can rule out the momentum-driven wind model at  $3\text{-}\sigma$  confidence level. We find no significant offset with respect to the fundamental metallicity relation, taking into account the intrinsic scatter and measurement uncertainties.

*Subject headings:* galaxies: abundances — galaxies: evolution — galaxies: formation — galaxies: high-redshift — gravitational lensing: strong

## 1. INTRODUCTION

Galaxies are complex ecosystems, particularly at the peak epoch of cosmic star formation, corresponding to the redshift range of  $z \sim 1\text{--}3$ , also known as the “cosmic noon” (see Madau & Dickinson 2014, for a recent review). During this approximately 4 Gyr, the Hubble sequence gradually breaks down and the predominant morphology of galaxies transforms from irregular systems at high redshifts to symmetric disks and bulges at low redshifts (Mortlock et al. 2013). This complex-

ity is to a large extent induced by the interplay between the process of star formation, and the diverse aspects of baryonic cycling, e.g., galactic feedback, gas inflows/outflows, and major/minor mergers (Davé et al. 2011b; Martin et al. 2012). The effect of environment surrounding galaxies can further complicate the spatial distribution of star formation, as recently revealed by Vulcani et al. (2015, 2016). Measurements of gas-phase metallicity, i.e., the chemical abundances of elements heavier than hydrogen and helium in the interstellar medium (ISM), are a powerful means to shed light on this

complexity<sup>1</sup>, because the metal enrichment history is strongly tied to the mass assembly history in galaxy evolution (Davé et al. 2011a; Lu et al. 2015b). Since detailed elemental abundances are not directly measurable at extragalactic distances, the relative oxygen abundance in ionized gaseous nebulae, i.e.,  $12 + \log(\text{O}/\text{H})$ , is often chosen as the observational proxy of metallicity.

For over a decade, a tight correlation between metallicity and galaxy stellar mass ( $M_*$ ), i.e., the mass-metallicity relation (MZR), has been quantitatively established, from the vast database of local galaxies observed by the Sloan Digital Sky Survey (Tremonti et al. 2004; Zahid et al. 2012; Andrews & Martini 2013). This relation has been further extended to high redshifts, using deep near infrared (IR) spectroscopy facilitated by large ground-based and space-based telescopes (Erb et al. 2006; Maiolino et al. 2008; Zahid et al. 2011; Henry et al. 2013; Steidel et al. 2014; Sanders et al. 2015; Guo et al. 2016). The measurements of the MZR as a function of redshift can cast useful constraints on various galaxy evolution models, since the slope of the MZR is sensitive to the properties of outflows, such as the mass loading factor and the outflow speed (see, e.g., Davé et al. 2012; Lu et al. 2015a). This slope can also be explained by variations of star-formation efficiency and gas mass fraction in galaxies with different stellar masses (see, e.g., Baldry et al. 2008; Zahid et al. 2014). The normalization of the MZR can shed light upon the stellar chemical yield across cosmic time (Finlator & Davé 2008).

It was first suggested by Mannucci et al. (2010) that there exists a so-called fundamental metallicity relation (FMR) in the 3D parameter space spanned by  $M_*$ , star-formation rate (SFR), and metallicity such that the MZR is merely a 2D projection of this more fundamental 3D manifold (see also Hunt et al. 2016a). This 3D scaling relation shows a tight scatter ( $\sim 0.05$  dex) in metallicity and is speculated to not evolve with  $z$ .

This proposed lack of evolution of the FMR is conjectured to be due to the interplay of two separate physical processes in control at different cosmic epochs.

- At low  $z$ , outflows of metal-enriched material are the predominant driver of the FMR. For local galaxies below  $10^{10} M_\odot$ , metallicity is in strong anti-correlation with SFR (Mannucci et al. 2010). This is attributed to the fact that more violent star formation leads to more powerful galactic feedback<sup>2</sup> and gas outflows, which are capable of removing the stellar nucleosynthesis yields from the galaxy, if the gravitational well is shallow. The different properties of the outflow (in particular the mass-loading factor) therefore set the slope of the MZR (Davé et al. 2012). However more massive galaxies are less prone to the loss of heavy elements from metal-loaded outflows because of the steeper gravitational potential wells that these outflows have to overcome. Thus more massive galaxies can more easily retain their metal production. This naturally explains the existence of the MZR and the observation that no correlation between metallicity and SFR is seen in galaxies with  $M_* \gtrsim 10^{10.9} M_\odot$  (Mannucci et al. 2010).

- At high  $z$ , inflows of pristine/metal-poor gas are key factor that governs the behavior of the FMR. The inverse correlation between metallicity and SFR is ascribed to the fact that the

infall of metal-poor gas lowers the global metallicity and triggers active star formation activities, because the gas reservoir in the galaxy is replenished (as suggested by, e.g., Hunt et al. 2016b).

In this context, the apparent redshift evolution of the MZR normalization originates primarily from sampling the FMR in terms of galaxies with different SFR. This picture is in accord with the gas regulator model proposed by Lilly et al. (2013), even though mergers can also play a subtle role in shaping the form of the FMR by increasing the scatter (Michel-Dansac et al. 2008). However, at high redshifts, the validity of the FMR is still under investigation (see, e.g., Sanders et al. 2015; Wuyts et al. 2014).

Spatially resolved chemical information provides a more powerful diagnostic tool about galaxy baryonic assembly than integrated metallicity measurements, especially at high redshifts. Because for non-interacting galaxies, their metallicity gradients are found to be highly sensitive to the properties of gas, i.e., the surface density, the existence of inflows/outflows, and the kinematic structure (Cresci et al. 2010; Jones et al. 2013; Sánchez et al. 2014; Troncoso et al. 2014). In the past few years, radial metallicity gradients, measured from spectroscopic data acquired by ground-based instruments with natural seeing ( $\gtrsim 0''.6$ ), have been reported at high redshifts (Queyrel et al. 2012; Troncoso et al. 2014; Stott et al. 2014). In particular, a large mass-selected sample of galaxies at  $0.7 \lesssim z \lesssim 2.7$  were observed with spatially resolved gas kinematics and star formation, using the K-band Multi Object Spectrograph (KMOS) on the Very Large Telescope (VLT) (i.e., the KMOS<sup>3D</sup> survey, Wisnioski et al. 2015). As a result, Wuyts et al. (2016) measured metallicity gradients for 180 star-forming galaxies in three redshift intervals (i.e. [0.8, 1.0], [1.3, 1.7], and [2.0, 2.6]), and found that the majority of the metallicity gradients are flat and no statistical significant correlations between metallicity gradients, and stellar, kinematic, and structural properties of the galaxy.

However, seeing-limited measurements typically have insufficient resolution to resolve the inner structure of galaxies at  $z \gtrsim 1$ , and the inferred metallicity gradient can be potentially biased. For example, Yuan et al. (2013) showed that coarse spatial sampling ( $\geq 1$  kpc) can result in artificially flat metallicity gradients, inferring that sufficiently high spatial resolution, i.e., *sub-kpc* scale (corresponding to an angular resolution of  $< 0''.2$ – $0''.3$ ), is crucial in yielding precise information of how metals are distributed spatially in extragalactic HII regions. Only a few metallicity gradient measurements meet this requirement, including a sample of 9 H $\alpha$ -selected galaxies at  $z \in [0.84, 2.23]$  (the majority at  $z \sim 1.45$ ), observed with the adaptive optics (AO) assisted integral field unit (IFU) spectrograph *SINFONI* onboard the VLT (Swinbank et al. 2012). Even higher resolution can be attained through the combination of diffraction-limited data and gravitational lensing as shown by Jones et al. (2010, 2013); Yuan et al. (2011), targeting strongly lensed galaxies with the laser guide star AO aided IFU spectrograph *OSIRIS* at the Keck telescope. Following this approach, Leethochawalit et al. (2015) recently analyzed a sample of 11 lensed galaxies (3 at  $z \sim 1.45$  and the rest at  $z \gtrsim 2$ ), deriving maps of both metallicity and emission line (EL) kinematics. Although a great amount of effort has been invested in enlarging the sample size of high- $z$  metallicity gradients obtained with sub-kpc scale spatial resolution, the current sample consists of only  $\lesssim 30$  such gradients and is still statistically insufficient to explore trends with stellar mass and redshift.

<sup>1</sup> Hereafter throughout the paper, we refer to gas-phase metallicity as metallicity for simplicity.

<sup>2</sup> Galactic feedback can suppress the star-formation efficiency by heating up the gas thereby preventing it from forming stars. So the interplay between star formation and feedback is self-regulating.

TABLE 1  
PROPERTIES OF THE GRISM SPECTROSCOPIC DATA USED IN THIS WORK

PA <sup>a</sup> (deg.)	Grism	Exposure Time (s)	Program	Time of completion
032	G102	8723	GLASS	February 2014
	G141	4412	GLASS	
111	G141	36088	SN Refsdal follow-up	December 2014
119	G141	36088	SN Refsdal follow-up	January 2015
125	G102	8623	GLASS	November 2014
	G141	4412	GLASS	

NOTE. — Here we only include the grism observations targetted on the prime field of MACS1149.6+2223.

<sup>a</sup> The position angle shown here corresponds to the “PA\_V3” value reported in the WFC3/IR image headers. The position angle of the dispersion axis of the grism spectra is given by  $PA_{\text{disp}} \approx PA_{\text{V3}} - 45^\circ$ .

In order to enlarge the sample of sub-kpc resolution measurements, we recently demonstrated that such metallicity maps can be derived using space-based data (Jones et al. 2015b). We measured a flat metallicity gradient in a multiply lensed interacting system at  $z = 1.85$ , using diffraction-limited *HST* grism data, and confirmed that flat metallicity gradients can be caused by gravitational interactions in merging systems.

The main goal of the work presented here is to collect a uniformly analyzed large sample of high- $z$  metallicity maps obtained with sub-kpc spatial resolution. To meet our goal, we improve upon the methodology proposed in our pilot paper (Jones et al. 2015b), in particular via developing a novel Bayesian method to imply metallicity from multiple EL diagnostics simultaneously. We apply our advanced analysis to ultra-deep grism data of the massive galaxy cluster MACS1149.6+2223, thus exploiting the powerful synergy of *HST* diffraction-limited spectroscopy and lensing magnification.

The outline of this paper is as follows. In Section 2, we introduce the spectroscopic observations used in this work and the selection of our metallicity gradient sample. The photometric data and galaxy cluster lens models are briefed in Section 3. We then present our entire analysis process in Section 4, and show our results in terms of global demographics and spatially resolved analysis in Sections 5 and 6, respectively. Finally, Section 7 will conclude and discuss our study. The concordance cosmological model ( $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ) and the AB magnitude system (Oke & Gunn 1983) are used throughout this paper. Without specific number showing the wavelength, the names of forbidden lines are simplified as [O II]  $\lambda\lambda 3726, 3729 :=$  [O II], [O III]  $\lambda 5008 :=$  [O III], [N II]  $\lambda 6585 :=$  [N II], [S II]  $\lambda\lambda 6718, 6732 :=$  [S II] in this paper<sup>3</sup>.

## 2. SAMPLE SELECTION, OBSERVATIONS, AND DATA REDUCTION

In Section 2.1, we first describe our *HST* grism observations, data quality, and data reduction. In Section 2.2, we then describe our sample selection criteria. In Section 2.3, we summarize the ground based kinematic data that we will use to supplement and interpret the metallicity and star formation maps obtained from the *HST* grism spectroscopy.

### 2.1. Hubble Space Telescope Grism Data

#### 2.1.1. The Grism Lens-Amplified Survey from Space

The Grism Lens-Amplified Survey from Space<sup>4</sup> (GLASS; Proposal ID 13459; P.I. Treu, Schmidt et al. 2014; Treu et al. 2015) is an *HST* cycle 21 large general observing (GO) program. GLASS observed 10 massive galaxy clusters with the Wide Field Camera 3 Infrared (WFC3/IR) grisms (G102 and G141; 10 and 4 orbits per cluster, respectively) targeted at their centers and the Advanced Camera for Survey (ACS) Optical grism (G800L) at their infall regions. The exposure on each cluster core is split into two nearly orthogonal position angles (PAs), in order to disentangle contamination from neighboring objects. Data acquisition was completed in January 2015. Here we focus on the grism data targeted on the center of MACS1149.6+2223 taken in February 2014 (PA=032°) and November 2014 (PA=125°), marked by blue squares in Figure 1 (also see Table 1). GLASS provides ~10% of the G141 exposures used in this work, and 100% of the G102 exposures. Details on GLASS data reduction can be found in Schmidt et al. (2014); Treu et al. (2015).

Shallow images through filters F105W or F140W were taken to aid the alignment and extraction of grism spectra. The imaging exposures are combined with the exposures obtained by the Hubble Frontier Fields (HFF) initiative and other programs to produce the deep stacks, released as part of the HFF program (see Section 3 for more details).

#### 2.1.2. SN Refsdal Follow-up Program with *HST* G141

The majority of the G141 exposures used in this work (30 orbits) were taken as part of the follow-up *HST* GO/DDT campaign (Proposal ID 14041, P.I. Kelly, Brammer et al. in prep.) of the first multiply imaged supernova, SN Refsdal (Kelly et al. 2015b). Two pointings (shown by the red squares in Figure 1) were exposed between December 2014 and January 2015 with 8 degrees apart (i.e. at PA=111°, and 119° respectively, as shown in Table 1), in order to optimize the spectroscopy of SN Refsdal. The analysis of the spectra of SN Refsdal, which showed that it was SN 1987A-like, is described by Kelly et al. (2015c).

#### 2.1.3. Grism data reduction

The combined grism dataset was reduced following the procedure of the 3D-*HST* survey (Brammer et al. 2012; Momcheva et al. 2015). An updated 3D-*HST* pipeline<sup>5</sup> was employed to reduce the images and spectra. The ASTRODRIZZLE software from the DRIZZLEPAC package and tweakreg were used to align and combine individual grism exposures, subtracting the sky images provided by Brammer et al. (2012). The time-varying sky background due to helium glow (at 10,830 Å) in the Earth exosphere was also accounted for according to Brammer et al. (2014). After these initial steps, the mosaics were co-added through interlacing onto a grid of 0.065 resolution, Nyquist sampling the point spread function (PSF) full width half maximum (FWHM). The average spectral resolution is  $R \sim 210$  and  $R \sim 130$  for G102 and G141, respectively. The average pixel scale after interlacing is 12 Å and 22 Å respectively for G102 and G141.

We used the co-added  $H_{160}$ -band mosaics as the detection image, on which SExtractor was run. An object catalog was generated and a corresponding segmentation map was created for each object, determining which spatial pixels (spaxels) belong to this object. Thus for a source registered in the catalog

<sup>3</sup> The wavelength values are taken from <http://classic.sdss.org/dr7/algorithms/linestable.html>

<sup>4</sup> <http://glass.astro.ucla.edu>

<sup>5</sup> <http://code.google.com/p/threedhst/>



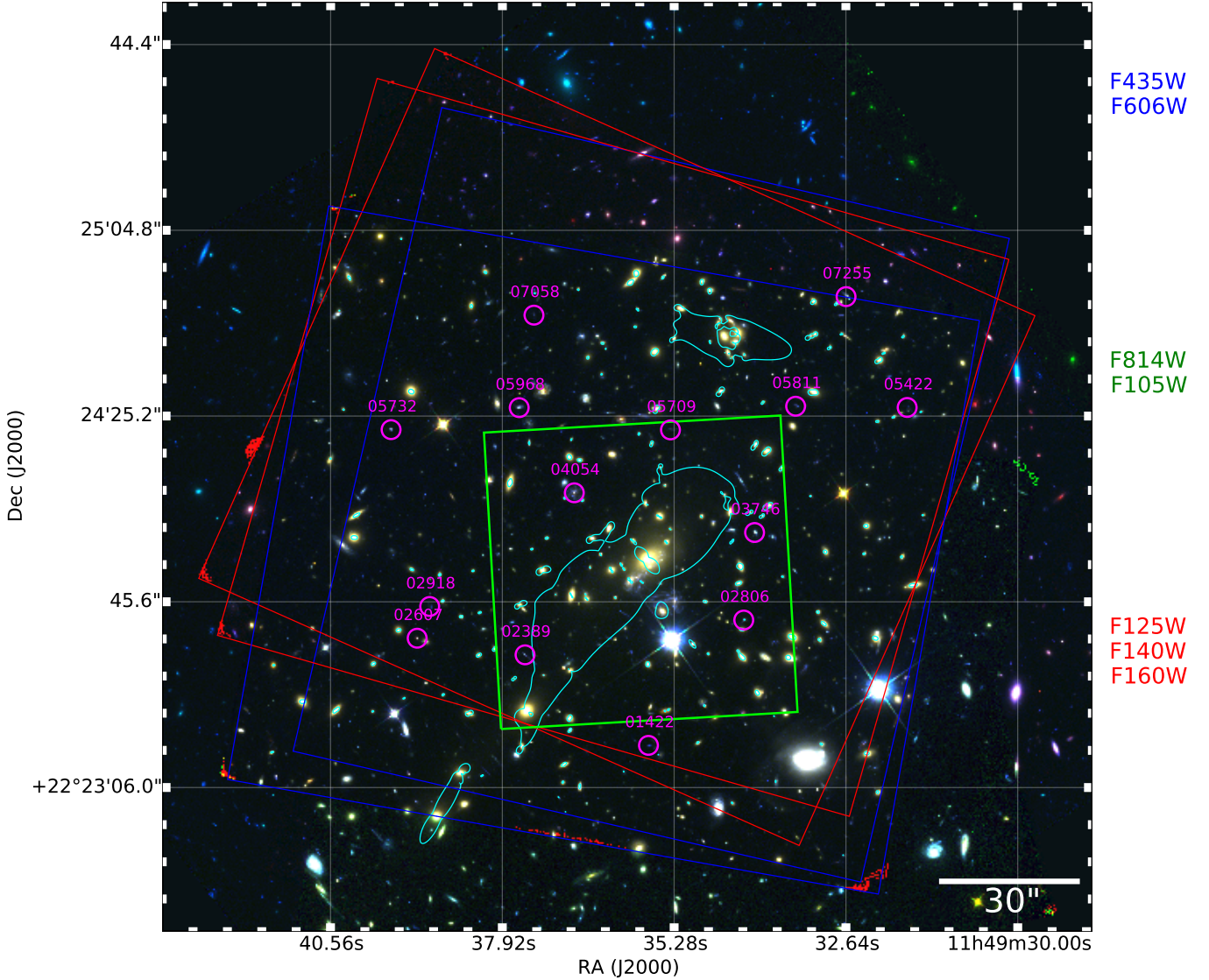


FIG. 1.—: Color composite image of MACS1149.6+2223 from the full-depth 7-filter *HFF* photometry. The blue, green, and red channels are comprised of the *HST* broad-band filters shown on the right. The blue, red and green squares mark the footprints of *GLASS*, the Refsdal follow-up, and *MUSE* programs utilized in this work. The cyan contours are the critical curves at  $z = 1.8$  predicted by the GLAFIC version 3 lens model. The magenta circles show the positions of our metallicity gradient sample.  $3 \times 3$  arcsec<sup>2</sup> zoom-in postage stamps around these positions are shown in Figure 2.

and falling within the WFC3 grism field-of-view (FoV), we extracted its spectra from grism mosaics through adding up the dispersed flux for all spaxels within an area defined by the segmentation image of this source, after flat-field correction and background subtraction. The “fluxcube” model from the AXE package (Kümmel et al. 2009) was employed to generate the 2D models of stellar spectrum, based upon the spatial profile of the source, the color information from direct image mosaics (if available), and the grism calibration configurations. For sources of interest, this model serves as the continuum model which we re-scaled and subtracted from the observed 2D spectra (see Section 4.3 for more details) in order to obtain pure EL maps. At the same time, if the object is a bright contaminating neighbour, this 2D model also functions as spectral contamination, which is subtracted before the continuum removal.

The data used in this work contain a total of 34 orbits of G141 and 10 orbits of G102, reaching a  $1\text{-}\sigma$  flux limit

(uncorrected for lensing magnification) of  $3.5(1.2) \times 10^{-18}$  erg s<sup>-1</sup> cm<sup>-2</sup> over the wavelength range of G102(G141), calculated from Schmidt et al. (2016). Our data provide an uninterrupted wavelength coverage in the range of 0.8–1.7  $\mu$ m and make this particular field one of the deepest fields probed by *HST* spectroscopy to date.

## 2.2. Sample selection

The selection of our metallicity gradient sample is based upon the master redshift catalog for MACS1149.6+2223<sup>6</sup>, published by the *GLASS* collaboration. As described by Treu et al. (2016), redshifts were determined by combining spectroscopic information from *GLASS* and the supernova (SN) Refsdal follow-up *HST* grism program, ground-based *MUSE* observations and Keck *DEIMOS* data.

For the purpose of this study, we compiled an exhaustive

<sup>6</sup> available at <https://archive.stsci.edu/prepds/glass/>



TABLE 2  
GLOBAL DEMOGRAPHIC PROPERTIES OF THE METALLICITY GRADIENT SAMPLE ANALYZED IN THIS WORK

ID	R.A. [deg.]	Dec. [deg.]	$z_{\text{spec}}$	$\mu^a$	$H_{160}$ magnitude [ABmag]	SED fitting		EL diagnostics		
						$\log(M_*/M_\odot)$	$A_V^S$	$12 + \log(\text{O}/\text{H})$	$\text{SFR}^b$ [ $M_\odot/\text{yr}$ ]	$A_V^N$
01422	177.398643	22.387499	2.28	2.35 [2.09, 2.26]	24.46	$9.29^{+0.07}_{-0.01}$	$< 0.01$	$8.26^{+0.11}_{-0.13}$	$15.10^{+16.76}_{-7.11}$	$< 1.07$
02389	177.406546	22.392860	1.89	43.25 [37.47, 50.28]	23.29	$8.43^{+0.06}_{-0.07}$	$< 0.01$	$7.88^{+0.16}_{-0.15}$	$< 6.37$	$1.36^{+0.69}_{-0.82}$
02607 <sup>c</sup>	177.413454	22.393843	1.86	3.01 [2.79, 2.98]	22.63	$10.25^{+0.06}_{-0.10}$	$1.70^{+0.42}_{-0.18}$	$7.70^{+0.13}_{-0.11}$	$> 36.51$	$2.26^{+0.31}_{-0.58}$
02806	177.392541	22.394921	1.50	4.82 [3.11, 3.51]	23.35	$8.72^{+0.19}_{-0.01}$	$0.50^{+0.01}_{-0.01}$	$7.96^{+0.15}_{-0.16}$	$1.96^{+1.37}_{-0.13}$	$< 0.18$
02918	177.412652	22.395723	1.78	2.95 [2.72, 2.91]	25.22	$8.37^{+0.15}_{-0.08}$	$0.10^{+0.34}_{-0.10}$	$8.42^{+0.11}_{-0.14}$	$< 70.32$	$1.25^{+0.72}_{-0.80}$
03746	177.391848	22.400105	1.25	3.78 [3.56, 3.83]	22.48	$8.81^{+0.03}_{-0.01}$	$< 0.01$	$8.11^{+0.10}_{-0.11}$	$10.95^{+1.61}_{-0.58}$	$< 0.17$
04054	177.403393	22.402456	1.49	3.39 [3.04, 3.28]	21.27	$9.64^{+0.05}_{-0.01}$	$1.10^{+0.01}_{-0.01}$	$8.70^{+0.09}_{-0.11}$	$16.99^{+5.71}_{-2.80}$	$0.72^{+0.23}_{-0.23}$
05422	177.382077	22.407510	1.97	3.63 [3.11, 4.41]	25.05	$8.45^{+0.15}_{-0.46}$	$0.00^{+0.56}_{-0.01}$	$7.56^{+0.16}_{-0.13}$	$< 38.18$	$< 1.85$
05709	177.397234	22.406181	1.68	7.10 [6.74, 7.53]	25.44	$7.90^{+0.02}_{-0.03}$	$< 0.01$	$8.21^{+0.12}_{-0.21}$	$< 17.97$	$< 1.18$
05732	177.415126	22.406195	1.68	1.56 [1.51, 1.61]	23.45	$9.09^{+0.01}_{-0.01}$	$< 0.01$	$8.41^{+0.10}_{-0.12}$	$< 84.70$	$< 1.31$
05811	177.389220	22.407583	2.31	7.50 [6.88, 9.34]	23.79	$8.85^{+0.11}_{-0.10}$	$< 0.01$	$8.16^{+0.14}_{-0.21}$	$4.94^{+3.13}_{-2.47}$	$< 0.74$
05968	177.406922	22.407499	1.48	1.84 [1.78, 1.96]	22.24	$9.34^{+0.02}_{-0.01}$	$0.60^{+0.20}_{-0.01}$	$8.47^{+0.07}_{-0.08}$	$27.95^{+4.70}_{-4.41}$	$0.23^{+0.15}_{-0.13}$
07058	177.405976	22.412977	1.79	2.13 [1.97, 3.01]	24.28	$8.74^{+0.17}_{-0.15}$	$0.00^{+0.12}_{-0.01}$	$8.38^{+0.13}_{-0.19}$	$< 44.99$	$< 1.85$
07255	177.385990	22.414074	1.27	2.34 [2.31, 2.99]	22.47	$9.36^{+0.01}_{-0.21}$	$0.30^{+0.21}_{-0.01}$	$8.38^{+0.08}_{-0.08}$	$8.73^{+3.53}_{-3.10}$	$< 0.79$

NOTE. — The error bars and upper/lower limits shown in the columns of SED fitting and EL diagnostics correspond to 1- $\sigma$  confidence ranges.

<sup>a</sup> Best-fit magnification values and 1- $\sigma$  confidence intervals. Except for galaxy ID 02389, the magnification results are from the GLAFIC version 3 model, calculated by the *HFF* interactive online magnification calculator available at <https://archive.stsci.edu/prepds/frontier/lensmodels/webtool/magnif.html>. For galaxy ID 02389, we use the SHARON & JOHNSON version 3 model instead to compute the magnification results and correct for lensing magnification.

<sup>b</sup> Values presented here are corrected for lensing magnification.

<sup>c</sup> The EL diagnostic result on this source is not trustworthy, since it is classified as an AGN candidate (see Sect. 5).

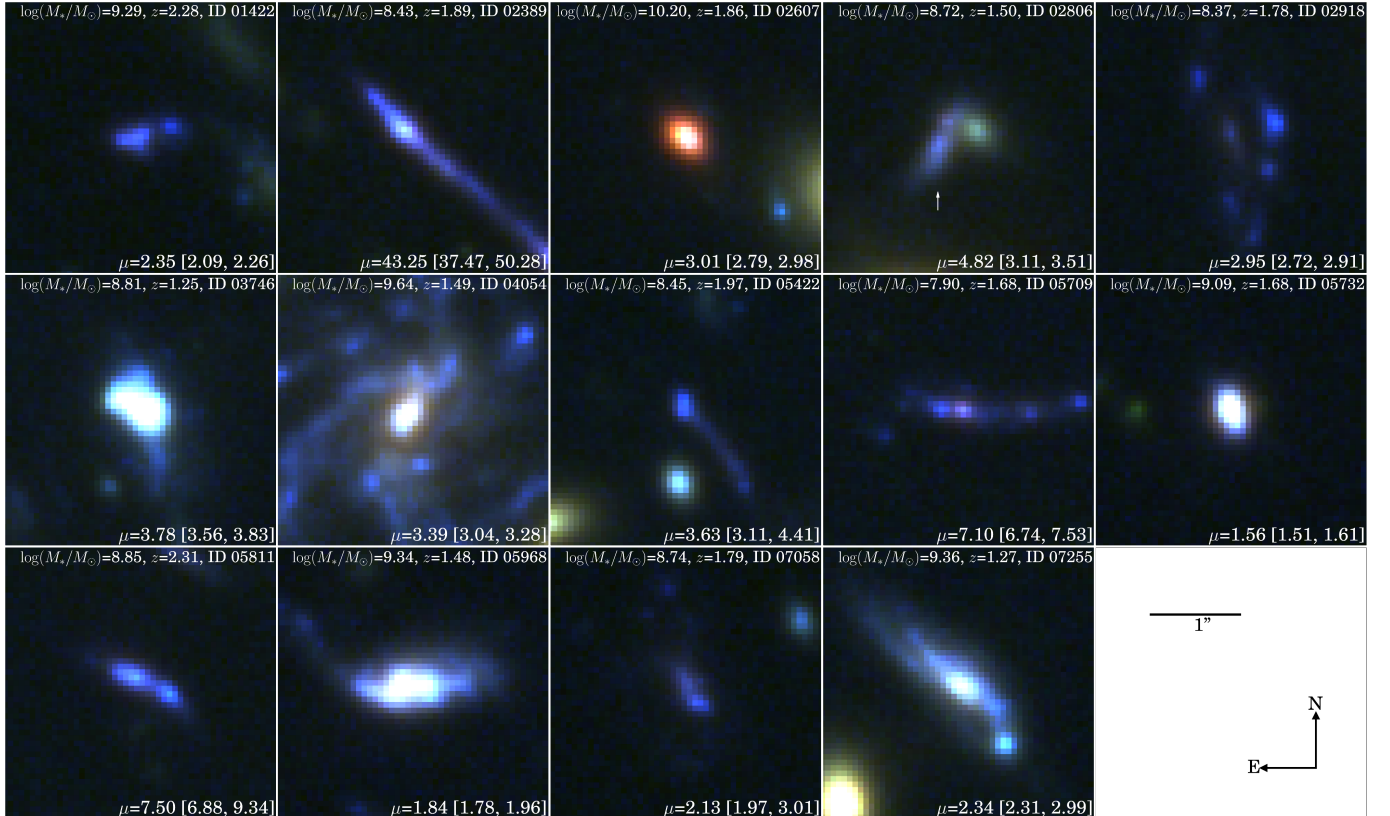


FIG. 2.— Zoom-in color composite stamps of our metallicity gradient sample, cut out from the full FoV image shown in Figure 1. In each stamp, we also show the spectroscopic redshift, stellar mass and lensing magnification (best-fit value followed by 1- $\sigma$  confidence interval) of the galaxy. In the panel of ID 02806, we use the white arrow to point to our galaxy of interest (ID 02806). All stamps are on the same spatial scale. The 1'' scale bar and north-east compass are shown in the lower right corner.

list of spatially extended sources with secure spectroscopic redshift measurements in the range of  $z \in [1.2, 2.3]$ , showing strong ELs (primarily [O III] and H $\beta$ ) in the 2D grism spectra. This redshift range avoids the region where grism sensitivity and throughput drop significantly. This results in a sample of 14 galaxies. The positions of these 14 galaxies relative to the cluster are denoted by the magenta circles in Figure 1, while the postage stamp images of these galaxy are shown in Figure 2. The full sample is also listed in Table 2. In particular, one source in our sample, i.e., galaxy ID 04054, is one of the multiple images of SN Refsdal’s host galaxy (see Figure 3 for its appearances from various perspectives). A steep metallicity gradient ( $-0.16 \pm 0.02$  dex/kpc) has been previously measured on another multiple image of this grand-design spiral (Yuan et al. 2011). Our work is the first metallicity gradient measurement on this particular image, which is least distorted/contaminated by foreground cluster members. We also take advantage of a more precise lens model of both the entire cluster and the SN Refsdal host, from the extensive lens modeling effort summarized by Treu et al. (2016).

### 2.3. Ground-based IFU observations

When available, we use ground based spectroscopic integral field unit (IFU) data to investigate the 2D kinematics of the sources for which we obtain star formation and metallicity maps from the *HST* grism data. A first set of ground-based data were obtained in 2015 with the instrument *MUSE* on the *VLT*, taken as part of the Director’s Discretionary Time program 294.A-5032 (P.I. Grillo) aimed at assisting with the modeling of SN Refsdal. The *MUSE* pointing is denoted by the green square in Figure 1. The observations, data reduction, first results, and applications to the host galaxy of Refsdal and mass modeling of the cluster are described by Grillo et al. (2016); Karman et al. (2016). Five objects in our metallicity gradient sample fall within the *MUSE* pointing, with one on the edge of the FoV. Hence four maps showing complete kinematic structures can be extracted for sources IDs 02389, 02806, 03746, and 04054. A second set of ground based observations were obtained with the instrument *KMOS* also on the *VLT*, as part of our *GLASS* follow-up *KMOS* Large Program 196.A-0778 (P.I. Fontana). Two objects (IDs 05709 and 05968) in our sample were targeted with the deployable IFUs, using the YJ grating. The observations, data reduction, and first results from this program are described by Mason et al. (2016).

### 3. PHOTOMETRY AND LENS MODELS FROM THE HFF

The Hubble Frontier Fields initiative<sup>7</sup> (*HFF*; P.I. Lotz, Lotz et al. 2016) is an ongoing multi-cycle treasury program enabled by an *HST* director’s discretionary time allocation. *HFF* targets the cores (prime fields) and infall regions (parallel fields) of six massive galaxy clusters, reaching an ultra-faint intrinsic magnitude of 30–33, made possible by the synergy of diffraction limited photometry and lensing magnification. The large wavelength coverage, uninterrupted from  $B_{435}$  to  $H_{160}$  passbands, is crucial for photometric redshift determinations and stellar population synthesis.

Besides the deep image mosaics, the *HFF* collaboration also provides the community with cluster lens models<sup>8</sup>. Several groups of scientists were invited prior to the beginning of the campaign to provide independent models depict-

ing the total mass distribution of the six *HFF* clusters, using a number of distinct techniques. As data accumulate and more multiply imaged systems are identified, the lens models are continuously improved. For our cluster of interest, MACS1149.6+2223, the most up-to-date publicly available model is the GLAFIC version 3 model, constructed using a simply parametrized method proposed by Oguri (2010). For an in-depth description of the various modeling techniques, their advantages and limitations, we refer to Treu et al. (2016); Meneghetti et al. (2016).

In this work, we need lens models to trace the observed metallicity maps back to the source plane. Following our previous work (Jones et al. 2015b), we experiment with several publicly available models to find the ones that give the reconstruction of our target sources the highest fidelity (judging by how well the source plane morphologies of multiply imaged sources match each other). Based upon the SN Refsdal test (see also Kelly et al. (2015a)), the GLAFIC version 3 model (Kawamata et al. 2015), the GRILLO model (Grillo et al. 2016), and the SHARON & JOHNSON version 3 model were the most accurate ones, so we considered those three. As an illustration, the GLAFIC version 3 critical curves at  $z = 1.8$ , the median redshift of our metallicity gradient sample, are overlaid in cyan in Figure 1.

The most highly magnified source in our sample is galaxy ID 02389, one of the folded arcs which straddle the critical curve. For that particular source, we used as our default the SHARON & JOHNSON version 3 model, updated from the earlier versions presented in Johnson et al. (2014); Sharon & Johnson (2015). The SHARON & JOHNSON version 3 model leads to a more precise reconstruction of the source plane morphology of galaxy ID 02389. We also tested that switching entirely from the GLAFIC version 3 model to the SHARON & JOHNSON version 3 model does not affect our measurements significantly, as also pointed out by Leethochawalit et al. (2015). The lensing magnification results from the considered lens models are given in Table 2.

### 4. ANALYSIS PROCEDURES

Here we describe the stellar population synthesis analysis of our sample in Section 4.1, and source plane morphology reconstruction in Section 4.2. The entire steps necessary for extracting EL maps from 2D grism spectra combined from different PAs are detailed in Section 4.3. Our new Bayesian method to jointly infer metallicity, nebular dust extinction, and star-formation rate from a simultaneous use of multiple strong ELs is presented in Section 4.4.

#### 4.1. Spectral energy distribution fitting

The full-depth 7-filter *HFF* photometry was fitted with the template spectra from Bruzual & Charlot (2003) using the stellar population synthesis code FAST (Kriek et al. 2009), in order to derive global estimates of stellar mass ( $M_*$ ) and the dust extinction of stellar continuum ( $A_V^S$ ) for each source in our sample. We take a grid of stellar population parameters that include: exponentially declining star formation histories with  $e$ -folding times ranging from 10 Myr to 10 Gyr, stellar ages ranging from 5 Myr to the age of the universe at the observed redshift, and  $A_V^S=0$ –4 magnitudes for a Calzetti et al. (2000) dust extinction law. We assume the Chabrier (2003) initial mass function (IMF) and fix the stellar metallicity to solar. Table 2 lists the best-fit  $M_*$  values, corrected for lensing magnification. Although the adopted solar metallicity is higher than the typical gas-phase abundances we infer for

<sup>7</sup> <http://www.stsci.edu/hst/campaigns/frontier-fields>

<sup>8</sup> <http://archive.stsci.edu/prepds/frontier/lensmodels/>

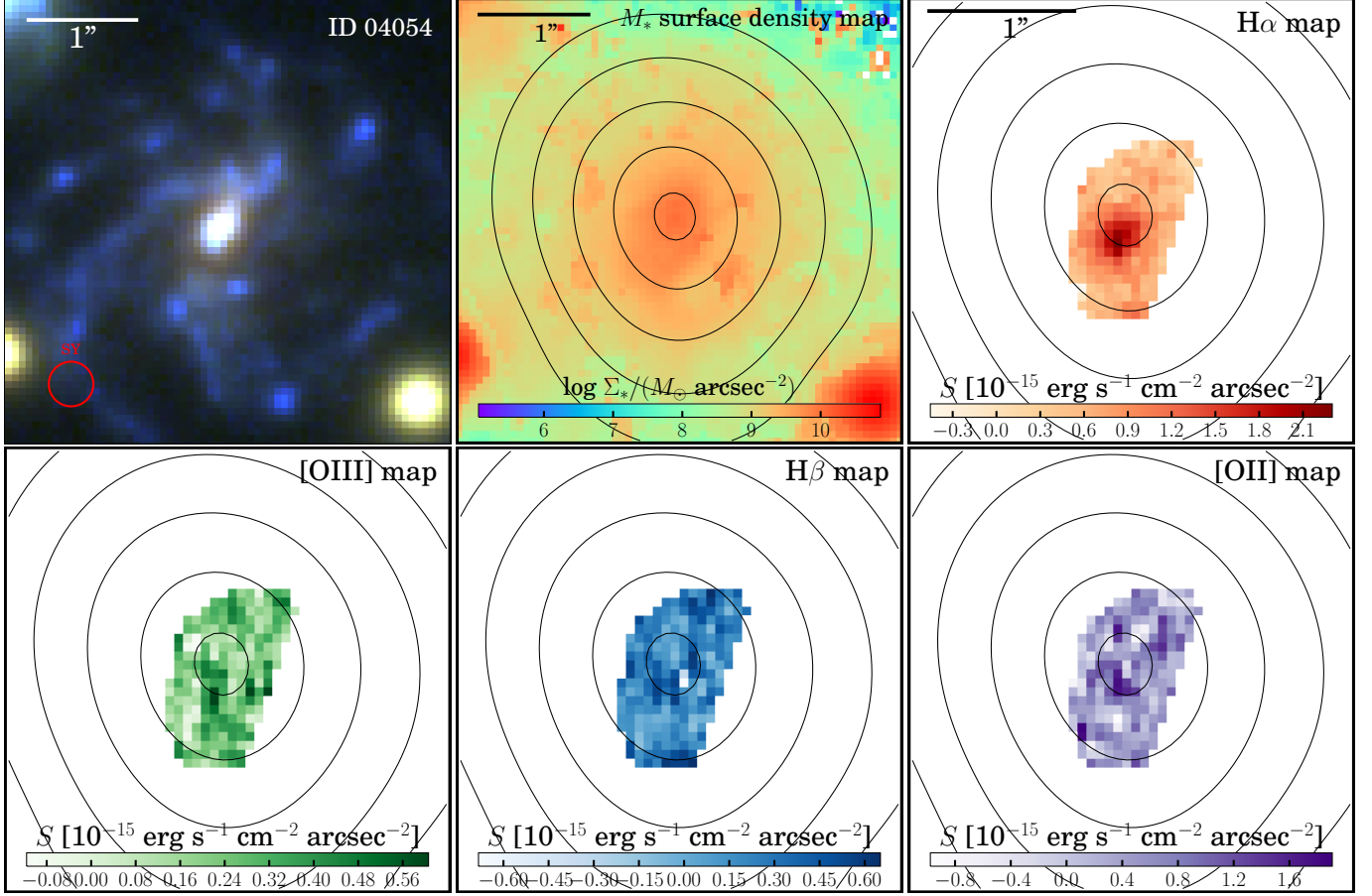


FIG. 3.— Multi-perspective view of the SN Refsdal host galaxy multiple image 1.3, i.e., ID 04054 in our metallicity gradient sample. **Upper left:** zoom-in color composite stamp cut out from Figure 1, where the estimated site of the un-observed past appearance of SN Refsdal is denoted by the red circle. **Upper middle:** stellar mass surface density map for this galaxy, where the source plane de-projected galactocentric radius contours are overlaid, in 2 kpc intervals, starting from 1 kpc. **Upper right:**  $H\alpha$  surface brightness map combined from the *HST* grism exposures at all PAs from the *GLASS* and Refsdal follow-up programs. The black contours again show the de-projected galactocentric radii in the same fashion presented in the stellar mass surface density map in the upper middle panel. **Lower rows:** combined surface brightness maps of [O III],  $H\beta$ , and [O II], presented in the same fashion as in the  $H\alpha$  emission line map in the upper right panel.

the sample, this has little effect on the derived stellar mass. Fixing the stellar metallicity to 1/5 solar (i.e.  $Z = 0.004$  or  $12 + \log(\text{O}/\text{H}) = 8.0$ , corresponding to the sample median) reduces the best-fit  $M_*$  by only  $\sim 0.03$  dex.

#### 4.2. Source plane morphology

Measurements of metallicity gradients require knowledge of the galaxy center, major axis orientation, and inclination or axis ratio. We derive these quantities from spatially resolved maps of the stellar mass surface density following the methodology described in Jones et al. (2015b). Briefly, we smooth the *HFF* photometric images to a common point spread function of  $0''.2$  FWHM and fit the spectral energy distribution (SED) of each spaxel using the same procedure described in Section 4.1. The resulting  $M_*$  maps in the image plane are shown in Figures 3 and 4. We reconstruct the  $M_*$  maps in the source plane using the adopted lens models and determine the centroid, axis ratio, inclination and major axis orientation from a 2D elliptical Gaussian fit to the  $M_*$  distribution. This allows us to measure the galactocentric radius at each point, assuming that contours of  $M_*$  surface density correspond to constant de-projected radius (following Jones et al. 2015b).

Since the following analysis is mostly done with the robust

image plane data, we reconstruct the derived radius maps in the image plane. Contours of constant galactocentric radius are shown in Figures 3 and 4. For half of our sample we also have seeing-limited IFU data described in Section 2.3, and verify that for galaxies with significant rotation support (e.g. IDs 05709 and 05968), the morphological major axis agrees at  $2\text{-}\sigma$  with the pseudo-slit that maximizes velocity shear according to our kinematic analysis (see Section 6.2).

#### 4.3. Emission line maps from grism spectra

The broad grism wavelength coverage provides spatially resolved maps of multiple ELs, such as [O II],  $H\gamma$ ,  $H\beta$ , [O III],  $H\alpha$  + [N II], and [S II]. To obtain pure EL maps, we first use the  $\lambda\text{Xe}$  continuum models described in Section 2 to subtract contaminating flux from neighboring sources. The continuum model for the target object is scaled to match the locally observed levels, and then subtracted. We check that the flux residual in regions near the ELs of interest follow a normal distribution with zero mean.

Due to the limited spectral resolution of *HST* WFC3/IR grisms,  $H\beta$  and the [O III]  $\lambda\lambda 4960, 5008$  doublets are partially blended in the spatially extended sources. We adopt a direct de-blending technique to separate these emission lines fol-



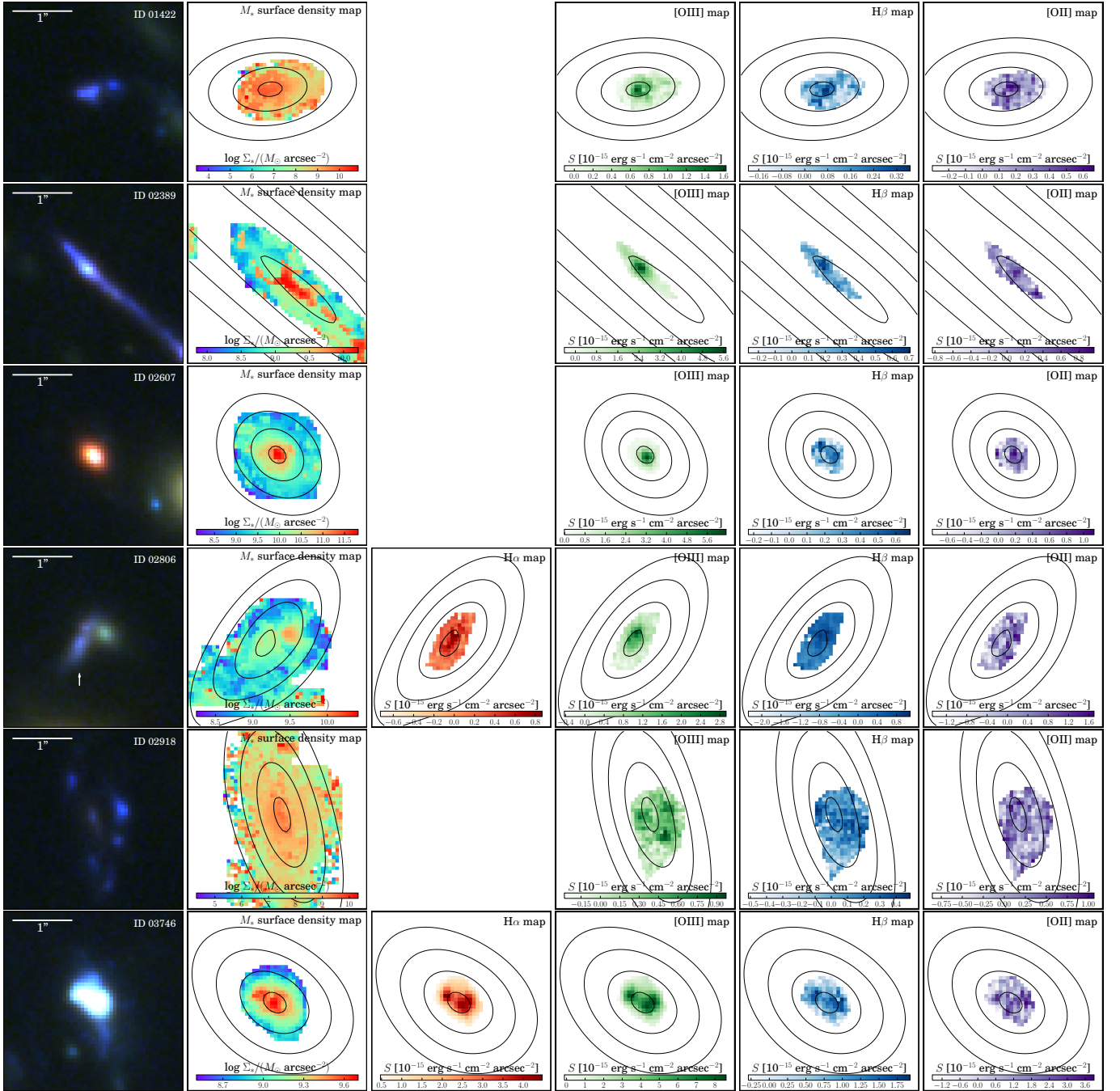


FIG. 4.—: Multi-perspective view of our metallicity gradient sample except for the SN Refsdal host which is shown in Figure 3. In each row, we show the zoom-in color composite stamp (the same as that in Figure 2), the stellar mass surface density map, the combined surface brightness maps of  $H\alpha$  (if available given object redshift), [O III],  $H\beta$ , [O II] for each object. The black contours overlaid represent the source plane de-projected galactocentric radii in 2 kpc interval, starting from 1 kpc. The spatial extent and orientation of the stamps for each object have been fixed.

lowing the procedure used in Jones et al. (2015b). First, an isophotal contour is measured from the co-added *HST*  $H_{160}$ -band postage stamp, typically at the level of  $\sim 10$  percent of the peak flux received from the source. We then apply this contour to the 2D grism spectra at the observed wavelengths corresponding to [O III]  $\lambda\lambda 4960, 5008$  and  $H\beta$ , maximizing the flux given grism redshift uncertainty ( $\sim 0.01$ ). We use the theoretical [O III]  $\lambda\lambda 4960, 5008$  doublet flux ratio (i.e. [O III]  $\lambda 5008$ /[O III]  $\lambda 4960 = 2.98$ , calculated by Storey & Zeippen (2000)) to subtract the flux contribution of [O III]  $\lambda 4960$  cor-

responding to the region where [O III]  $\lambda 5008$  is unblended. We iterate this procedure to remove the [O III]  $\lambda 4960$  emission completely, resulting in pure maps of  $H\beta$  and [O III]  $\lambda 5008$ .

Note that this direct de-blending will be compromised if the source is so extended that its [O III]  $\lambda 5008$  and  $H\beta$  are blended. In practice, we fine-tune the isophotal contour level to avoid this contamination. As a result, no cases in our sample have problem with our direct de-blending method, and the resulting pure [O III]  $\lambda 5008$  and  $H\beta$  maps are uncontaminated.

In the case of  $H\alpha$  and [N II], the emission lines are sep-

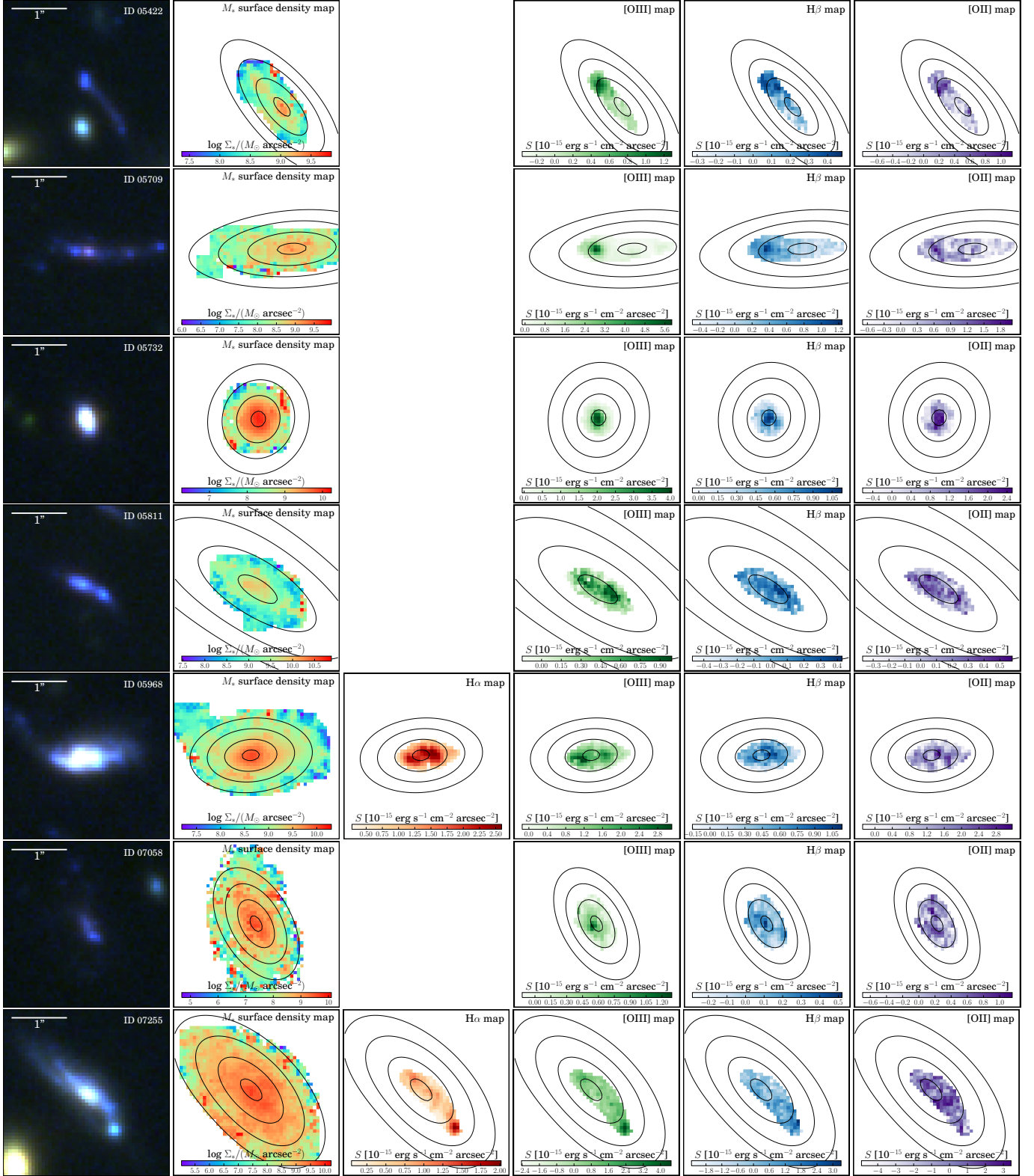


FIG. 4.— (cont.)

TABLE 3  
SAMPLING PARAMETERS AND THEIR PRIOR INFORMATION.

Set	Parameter	Symbol (Unit)	Prior
Vanilla	gas-phase metallicity	$12 + \log(\text{O}/\text{H})$	flat: [7.0, 9.3]
	nebular dust extinction	$A_V^N$	flat: [0, 4]
	intrinsic $\text{H}\beta$ flux	$f_{\text{H}\beta}^t$ ( $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ )	Jeffrey's: [0, 100]
Extended	[N II] to $\text{H}\alpha$ flux ratio	[N II]/ $\text{H}\alpha$	Jeffrey's: [0, 0.5]
Derived	star-formation rate	SFR ( $M_\odot/\text{yr}$ )	—

NOTE. — Most constraints presented in this work are obtained under the “Vanilla” parameter set with the extended parameter fixed as [N II]/ $\text{H}\alpha$  = 0.05.

arated by less than the grism spectral resolution. Therefore we cannot use the direct technique to de-blend  $\text{H}\alpha$  and [N II] emission, and instead treat the [N II]/ $\text{H}\alpha$  ratio as an unknown parameter in Section 4.4.

After obtaining pure EL maps, we combine data taken at multiple PAs in order to utilize the full depth of the grism exposures. There are in total six different PA-grism combinations as given in Table 1. At certain redshifts, ELs are visible in the overlapping region between the two grisms at  $\lambda \in [1.08, 1.15] \mu\text{m}$  resulting in up to six individual maps of the same EL. We use the data taken at PA=119° as a reference, as this minimizes errors arising from rotating the various PAs into alignment. We first apply a small vertical offset to each EL map to account for the slight tilt of the spectral trace (i.e. known offset between the dispersion direction and the x-axis of extracted spectra; Brammer et al. 2012). Next, we apply a small horizontal shift to correct for uncertainties in the wavelength calibration and redshift. Finally we rotate each EL map to align it with the reference orientation at PA=119°. The best values of these three alignment parameters are found by maximizing the normalized cross-correlation coefficient (NCCF) in order to achieve optimal alignment of multiple PAs:

$$\text{NCCF} = \frac{1}{n} \sum_{x,y} \frac{(S_{\text{ref}}(x,y) - \langle S_{\text{ref}} \rangle)(S_{\text{PA}}(x,y) - \langle S_{\text{PA}} \rangle)}{\sigma_{S_{\text{ref}}} \sigma_{S_{\text{PA}}}}. \quad (1)$$

Here  $S$  is the surface brightness profile for a specific EL,  $S_{\text{ref}}$  corresponds to the reference alignment stamp adopted as the PA=119° data, and  $n$  is the number of spaxels in the surface brightness profile.  $\langle S \rangle$  and  $\sigma_S$  represent the average and standard deviation of the surface brightness, respectively.

Once the data from each PA are aligned, we vet the EL maps from each PA-grism combination and reject those which show significant contamination-subtraction residuals or grism reduction defects. We combine the remaining maps of a given EL using an inverse variance weighted average in order to properly account for the different exposure times and noise levels. The final combined maps of each EL for each galaxy are shown in Figures 3 and 4.

#### 4.4. Line-flux-based Bayesian inference of metallicity

Generally speaking, two methods exist in deriving gas-phase oxygen abundance in star-forming galaxies at high redshifts. “Direct” determinations based on electron temperature measurements, which require high signal-to-noise ratio (SNR) detections of auroral lines (e.g. [O III]  $\lambda 4364$ ), are still extremely challenging beyond the local universe (see Sanders et al. 2016, for a rare example). We therefore rely on the calibrations of strong collisionally excited EL flux ratios to estimate metallicities.

One of the most popular strong EL diagnostics is the flux ratio of [N II]/ $\text{H}\alpha$  calibrated by Pettini & Pagel (2004). How-

ever, a large offset (0.2-0.4 dex) between the loci of high- $z$  and present-day star-forming galaxies in the Baldwin-Phillips-Terlevich (BPT, Baldwin et al. 1981) diagram has recently been discovered, indicating that extending the locally calibrated [N II]/ $\text{H}\alpha$  to high- $z$  can be problematic (Shapley et al. 2015; Sanders et al. 2015). The interpretation of this offset is still the topic of much debate. It has been interpreted as the existence of a fundamental relation between nitrogen-oxygen abundance ratio and  $M_*$  (Masters et al. 2016), a systematic dependence of strong EL properties with respect to Balmer line luminosity (Cowie et al. 2016), a combination of harder ionizing radiation and higher ionization states of the gas at high- $z$  (Steidel et al. 2014), or an enhancement of nitrogen abundance in hot H II regions (Pilyugin et al. 2010). Fortunately, the alpha-element BPT diagrams show no offset with  $z$  (Shapley et al. 2015; Sanders et al. 2015), and therefore the diagnostics based upon those ELs are more reliable.

In this work we adopt the calibrations of Maiolino et al. (2008, M08), including the flux ratios of [O III]/ $\text{H}\beta$ , [O II]/ $\text{H}\beta$ , [N II]/ $\text{H}\alpha$ . Given the potential systematics related to nitrogen ELs, we do not use the [N II]/ $\text{H}\alpha$  calibration of M08. M08 fit the relation between these EL ratios and gas-phase oxygen abundance based upon “direct” measurements (from [O III]  $\lambda 4364$ ) for  $12 + \log(\text{O}/\text{H}) < 8.3$ , and photoionization model results for higher metallicities. This provides a continuous framework valid over the wide range  $12 + \log(\text{O}/\text{H}) \in (7.1, 9.2)$ . We note that although there are alternative calibrations of metallicity available in the literature and the applicability of these recipes is currently a hotly debated topic (see, e.g., Blanc et al. 2015; Dopita et al. 2013). Here we are primarily interested in relative metallicity measurements. Even though the absolute measurements of metallicity may change if we used a different calibration, the gradients and morphological features in the maps should be more robust to changes in the calibration. We will restrict our comparisons with other samples to only include studies assuming the M08 calibrations.

Unlike previous work in which calibrated relations are applied to various EL flux ratios (see, e.g., Pérez-Montero 2014; Bianco et al. 2015), we design a Bayesian statistical approach which uses directly the individual EL fluxes as input, such that the information from one EL is only used once. Our approach presents several advantages over those based on line flux ratios. First, it properly accounts for flux uncertainties for lines that are weak or undetected. Second, multiple lines can be taken into consideration without the risk of double counting information.

Our approach simultaneously infers the gas-phase metallicity ( $12 + \log(\text{O}/\text{H})$ ), nebular dust extinction ( $A_V^N$ ), and intrinsic  $\text{H}\beta$  flux (which is proportional to the star formation rate). For values of these three parameters we can compute the expected line fluxes, and compare with measured values to compute the likelihood and then the posterior probability.

Expected line fluxes are uniquely determined from the M08 calibrations and the Cardelli et al. (1989) galactic extinction curve with  $R_V = 3.1$ , assuming intrinsic  $\text{H}\alpha/\text{H}\beta = 2.86$  and  $\text{H}\gamma/\text{H}\beta = 0.47$  appropriate for case B recombination and fiducial H II region conditions (see, e.g., Hummer & Storey 1987). Together with measured EL fluxes and uncertainties, the  $\chi^2$  statistic in our inference procedure is calculated as

$$\chi^2 = \sum_i \frac{(f_{\text{EL}_i}^o - R_i \cdot f_{\text{H}\beta}^t)^2}{(\sigma_{\text{EL}_i}^o)^2 + (f_{\text{H}\beta}^t)^2 \cdot (\sigma_{R_i})^2}. \quad (2)$$



Here  $EL_i$  denotes the set of emission lines [O II], H $\gamma$ , H $\beta$ , [O III], H $\alpha$ , and [S II].  $R_i$  is the flux of  $EL_i$  relative to H $\beta$ , with  $\sigma_{R_i}$  being intrinsic scatter from the M08 calibrations.  $f_{EL_i}^o$  and  $\sigma_{EL_i}^o$  represent the observed  $EL_i$  flux and its uncertainty.  $f_{H\beta}^t$  is the intrinsic H $\beta$  flux corrected for dust extinction.

We apply this inference procedure to both galaxy-integrated line fluxes (to be discussed in Section 5) as well as individual spaxels in the EL maps (which will lead to metallicity gradient measurements described in Section 6).

Details of the prior assumptions applied to each parameter of the inference procedure are explained in Table 3, and the coefficients for the EL flux ratio diagnostics used in the statistical inference are given in Table 4. We additionally note the following:

- The [N II]/H $\alpha$  ratio is included as an additional parameter for galaxies at  $z \lesssim 1.5$  where these lines are observed. We do not attempt to determine [N II]/H $\alpha$  using the M08 calibrations, as locally-calibrated diagnostics tend to underestimate the [N II]/H $\alpha$  ratio in high- $z$  galaxies whereas oxygen and other  $\alpha$ -element line ratios remain constant with redshift (Steidel et al. 2014; Shapley et al. 2015; Jones et al. 2015a). Instead we either leave this parameter free (“extended” priors), or fixed to [N II]/H $\alpha$  = 0.05 (“vanilla” priors). The vanilla value is typical of galaxies with similar  $z$  and oxygen EL ratios as our sample. Fixing the value of [N II]/H $\alpha$  provides faster convergence and does not significantly affect the inferred metallicity. We therefore report constraints for the “vanilla” parameter set. For galaxy ID 04054, the SN Refsdal host galaxy, we fix its [N II]/H $\alpha$  to be 0.11, as measured by Yuan et al. (2011).

- For  $EL_i \in \{H\alpha, H\gamma\}$ ,  $R_i$  corresponds to the Balmer decrement and we set  $\sigma_{R_i} = 0$  (i.e. the intrinsic Balmer ratios are fixed with no assumed scatter). These ratios are sensitive only to the nebular dust extinction  $A_V^N$ .

- For  $EL_i \in \{[O II], [O III]\}$ ,  $R_i$  and  $\sigma_{R_i}$  are taken from the M08 calibrations. These ratios are sensitive to the oxygen abundance  $12 + \log(O/H)$ , whereas dust extinction correction is necessary for [O II]/H $\beta$ .

- For  $EL_i = [S II]$ ,  $R_i$  and  $\sigma_{R_i}$  come from our new calibration derived following Jones et al. (2015a), using the same data as the low-metallicity calibrations of M08. Although M08 do not provide any calibrations for [S II], we expect our calibration to be self-consistent and reliable for  $12 + \log(O/H) \lesssim 8.4$ . [S II]/H $\alpha$  is primarily useful as a diagnostic to differentiate between the upper and lower branches of the [O III]/H $\beta$  calibration. This is valuable because [N II] is not directly measured and [O II] is typically low signal-to-noise in the lower redshift galaxies for which [S II] is available, due to decreasing grism throughput at  $\lambda < 0.9 \mu m$ .

- For  $EL_i = H\beta$ ,  $R_i = 1$  with  $\sigma_{R_i} = 0$ .

We use the Markov Chain Monte Carlo sampler EMCEE (Foreman-Mackey et al. 2013) to explore the parameter space, setting the number of walkers to 100 with 5,000 iterations for each walker, and a burnin of 2,000. The autocorrelation times are computed for each parameter and are used to make sure chains have converged. An example constraint result for the SN Refsdal host galaxy’s global EL fluxes is shown in Figure 5. While this paper is primarily concerned with gas-phase metallicities, our method simultaneously provides constraints on  $A_V^N$  and star-formation rate which can be compared to the results of SED fitting. We then derive the instantaneous star-formation rate from the intrinsic H $\alpha$  luminosity calcu-

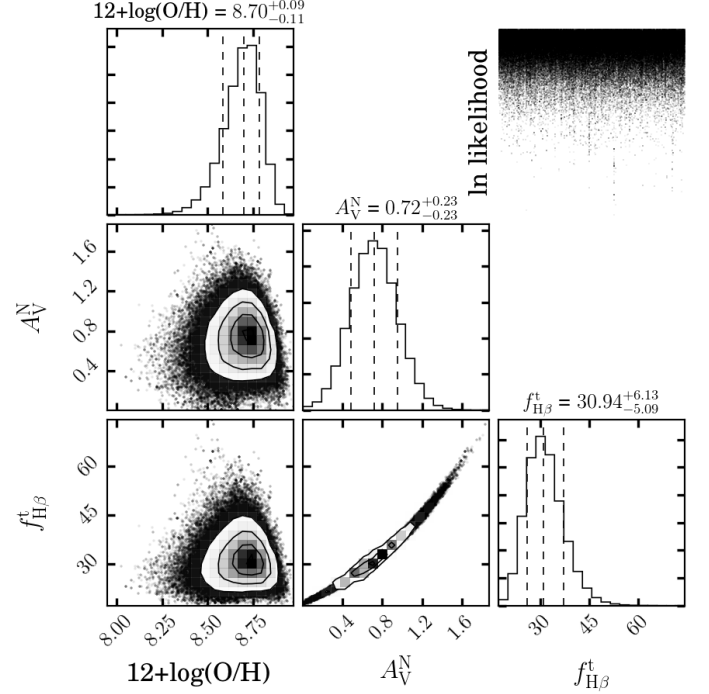


FIG. 5.—: The marginalized 1D and 2D constraints on  $(12 + \log(O/H), A_V^N, \text{ and } f_{H\beta}^t)$ , which are drawn from the integrated line fluxes of the Refsdal host galaxy (ID 04054), after all grism exposures are combined. The values on top of each column are the medians with  $1-\sigma$  uncertainties for each parameter. The plot in the upper right corner illustrates a good convergence of the sampling.

TABLE 4  
COEFFICIENTS FOR THE EL FLUX RATIO DIAGNOSTICS USED IN OUR BAYESIAN INFERENCE.

$R$	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$
H $\alpha$ /H $\beta$	0.4564	...	...	...	...	...
H $\gamma$ /H $\beta$	-0.3279	...	...	...	...	...
[O III]/H $\beta$	0.1549	-1.5031	-0.9790	-0.0297	...	...
[O II]/H $\beta$	0.5603	0.0450	-1.8017	-1.8434	-0.6549	...
[S II]/H $\alpha$	-0.5457	0.4573	-0.8227	-0.0284	0.5940	0.3426

NOTE. — The value of EL flux ratio is calculated by the polynomial functional form, i.e.,  $\log R = \sum_{i=0}^5 c_i \cdot x^i$ , where  $x = 12 + \log(O/H) - 8.69$ . The ratio of [S II]/H $\alpha$  is calibrated by the work of Jones et al. (2015a), and the Balmer line ratios correspond to the Balmer decrement, whereas the ratios between oxygen lines and H $\beta$  are from M08.

lated from  $f_{H\beta}^t$  via the commonly-used calibration of Kennicutt (1998a); Moustakas et al. (2010), i.e.,

$$\text{SFR} = 4.6 \times 10^{-42} \frac{L(H\alpha)}{\text{erg/s}} [M_{\odot}/\text{yr}], \quad (3)$$

valid for a Chabrier (2003) IMF. This SFR estimate does not rely on any assumptions about star-formation history, and thus is a more direct measure of ongoing star formation than values given by SED fitting.

## 5. GLOBAL DEMOGRAPHIC PROPERTIES

Based upon the Bayesian analysis, we obtain the global measurements of gas-phase metallicity, nebular dust extinction, and star-formation rate, for all sources in our sample,

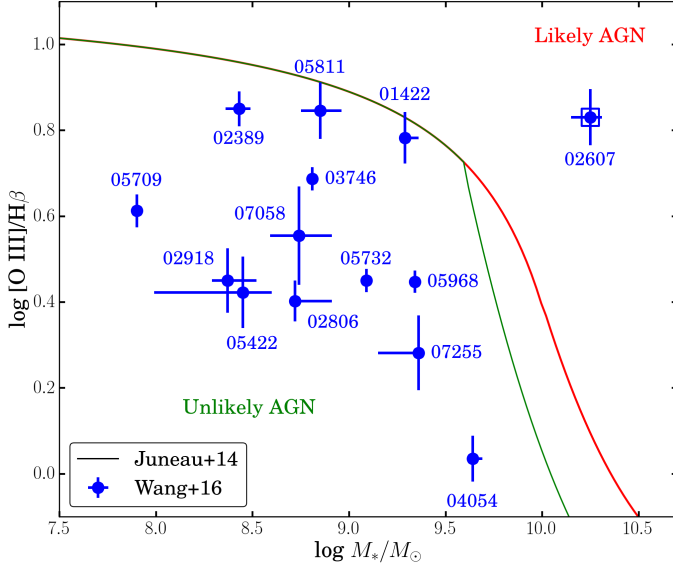


FIG. 6.—: Mass excitation diagnostic diagram for our metallicity gradient sample. The curves are the demarcation lines updated by Juneau et al. (2014). The regions below the green curve have a low probability ( $< 10\%$ ) of being classified as AGNs, whereas galaxies located above the red curve are secure AGN candidates. In our sample, the only AGN candidate, i.e., ID 02607, is marked by a square.

as shown in Table 2. First, we check for the presence of any active galactic nucleus (AGN) contamination in our sample. Because not all galaxies in our sample reside in the redshift range where all BPT lines are available, we resort to the mass-excitation diagram showing the  $[\text{O III}]/\text{H}\beta$  flux ratio as a function of  $M_*$ , first proposed by Juneau et al. (2011) to differentiate star-forming galaxies and AGNs in local Sloan Digital Sky Survey (SDSS) observations. Juneau et al. (2014) further revised the demarcation scheme by applying line luminosity evolution and selection effects to a more complete SDSS galaxy sample, and tested that this scheme agrees well with the bivariate distributions found in a number of high- $z$  galaxy samples. According to the Juneau et al. (2014) classification scheme, all but one galaxy (ID 02607) have a negligible probability of being AGN (see Figure 6). Recently, Coil et al. (2015) found that a  $+0.75$  dex  $M_*$  shift in the Juneau et al. (2014) scheme is required to avoid serious AGN contamination in the *MOSDEF* galaxy sample at  $z \sim 2.3$ . We verified that galaxy 02607 can still be safely classified as an AGN candidate, even considering the Coil et al. (2015) shifted classification scheme. Because the M08 calibrations originate from HII region observations and theoretical models, not designed for AGNs, our method will not predict a reliable metallicity constraint for the AGN candidate ID 02607. We thereby exclude this source in the rest of our population analysis. For the rest of our sample, comprising 13 star-forming galaxies, the constraint on gas-phase metallicity is robust.

In Figure 7, we plot the measured SFR as a function of  $M_*$  for our sample. In comparison, the loci of mass-selected galaxies observed by the *KMOS*<sup>3D</sup> survey and the “star-formation main sequence” (SFMS, Whitaker et al. 2014; Speagle et al. 2014) at similar redshifts are also shown. We can see that selecting lensed galaxies based upon EL strength can probe deep into the low-mass regime at high- $z$  that is currently inaccessible to mass complete surveys. As expected,

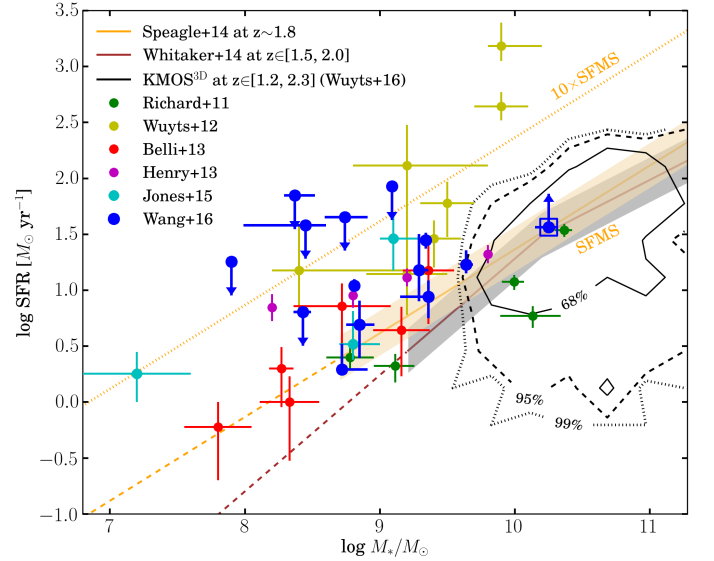


FIG. 7.—: Star formation rate as a function of stellar mass for emission-line galaxies at  $z \in [1.2, 2.3]$ . Our metallicity gradient sample is marked by blue points; the literature sample is color-coded by reference according to the legend. The star-forming main sequence compiled by Speagle et al. (2014) and Whitaker et al. (2014) are represented by brown and orange lines respectively, where the dotted part is extrapolated at masses below the mass completeness limit. The shaded regions denote the typical scatter of star-forming main sequence, i.e. 0.2-0.3 dex. We also show the source density contours for the *KMOS*<sup>3D</sup> survey at the same redshift range, where 68%, 95%, and 99% of all *KMOS*<sup>3D</sup> galaxies at  $z \in [1.2, 2.3]$  can be found within the black solid, dashed, and dotted contours respectively. It is found that our sample is highly complementary to the *KMOS*<sup>3D</sup> sample in terms of stellar mass.

our emission line selected targets probe the upper envelope of the SFMS, so that a detail comparison is non-trivial and should take into account the selection functions of each sample. The advantage of the emission line selection technique is that it is the most cost-effective way to obtain gas metallicities for stellar masses down to  $10^7 M_*$  at  $z \sim 2$ . Using the same technique, we discovered an analog of local group dwarf spheroidals (e.g., Fornax, Coleman & de Jong 2008) at  $z = 1.85$  experiencing active star formation (with 1 dex offset from the SFMS), in our previous work (Jones et al. 2015b).

We also collect all published gas-phase metallicity measurements in emission-line galaxies in the redshift range of our sample  $1.2 \lesssim z \lesssim 2.3$ , obtained exclusively from the strong EL calibrations prescribed by M08. The reason we only select measurements based upon the M08 calibrations is that adopting different strong EL calibrations can give rise to different absolute metallicity scales offset by up to  $\sim 0.7$  dex at the high-mass end, according to the comparative study conducted by Kewley & Ellison (2008). In order to minimize the systematic uncertainty associated with EL calibrations, we thus focus on this homogeneous M08 sample alone.

As shown in Figure 8, the M08-based sample includes five galaxies from Richard et al. (2011), seven from Wuyts et al. (2012), six from Belli et al. (2013), four stack measurements by Henry et al. (2013), three interacting galaxies at  $z = 1.85$  analyzed by Jones et al. (2015b), and 13 star-forming galaxies presented in this work. In total, this sample consists of 38

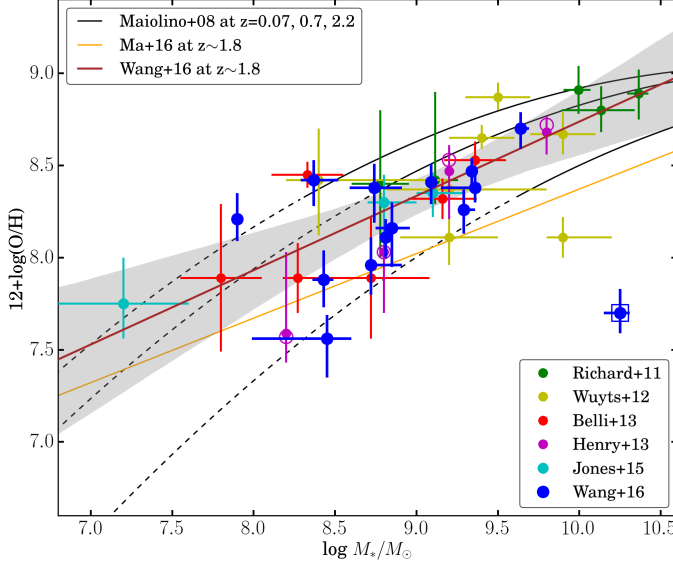


FIG. 8.— The relation between stellar mass and gas-phase oxygen abundance, from observations and simulations at  $z \sim 1.8$ . The color-coding for all the points is the same as in Figure 7. For the sake of consistency, all metallicity measurements are derived assuming the Maiolino et al. (2008) strong EL calibrations. The black curves represent the 2nd order polynomials fitted by Maiolino et al. (2008) to data sets at different redshifts, where dashed parts correspond to extensions beyond mass completeness limits of those data sets. The thick brown line denotes our best-fit linear relation based upon all data points except the AGN candidate (i.e. ID 02607) marked by a square. The shaded band marks the  $1-\sigma$  confidence region. For the Henry et al. (2013) stack data points, we also show the measurements without dust correction as open circles, which are not included in the fit. The thin orange line is the prediction from the cosmological zoom-in simulations conducted by Ma et al. (2016a) at the same redshift range, which is consistent with our measurement at  $1-\sigma$ .

independent measurements at  $1.2 \lesssim z \lesssim 2.3$ , with median redshift  $z_{\text{median}} = 1.84$ . The mutual agreement between our sample and that from the literature provides an independent confirmation that our new Bayesian method leads to constraints on gas-phase metallicity consistent with previous studies also adopting the same calibrations. With the combined sample we are able to derive the MZR at this redshift, spanning three orders of magnitude in stellar mass.

We fit the following functional form to this sample of 38 galaxies in order to derive the MZR, i.e.,

$$12 + \log(\text{O}/\text{H}) = \alpha + \beta \log(M_*/M_\odot) + N(\sigma), \quad (4)$$

where  $\alpha$ ,  $\beta$ , and  $\sigma$  represent the intercept, the slope and the intrinsic scatter, respectively. The Python package `linmix`<sup>9</sup> was employed to perform a Bayesian linear regression following the method proposed by Kelly (2007). The median values and  $1-\sigma$  uncertainties for these three parameters drawn from the marginalized posteriors are

$$\alpha = 4.71^{+0.62}_{-0.64}, \beta = 0.40^{+0.07}_{-0.07}, \sigma = 0.03^{+0.02}_{-0.01}. \quad (5)$$

The resulting MZR together with its  $1-\sigma$  confidence region is shown in Figure 8.

Recently, Guo et al. (2016) presented a comprehensive study of the MZR and its scatter at  $z=0.5-0.7$  from data acquired by large spectroscopic surveys in the CANDELS fields (Grogin et al. 2011; Koekemoer et al. 2011). Guo et al. (2016) also summarized a variety of theoretical predictions of the MZR slope predicted by diverse approaches, including those predicted by the energy-driven wind ( $\sim 0.33$ ) and the momentum-driven wind ( $\sim 0.17$ ) equilibrium models proposed by Finlator & Davé (2008); Davé et al. (2012). Our inferred slope is only marginally compatible with the prediction given by the energy-driven wind model at  $1-\sigma$  confidence level, whereas strongly disfavors the momentum-driven wind model. Considering the majority of the galaxies in our sample have relatively lower masses, our result confirms the finding by Henry et al. (2013) that momentum-driven winds cannot be the primary workhorse that shapes the MZR in the low-mass regime (below  $10^{9.5} M_\odot$ ) at  $z \gtrsim 1$ .

Interestingly, our result is consistent within  $1-\sigma$  with the formula of the MZR evaluated at  $z = z_{\text{median}}$ , i.e.,  $12 + \log(\text{O}/\text{H}) = 0.35 \log(M_*/M_\odot) + 4.87$ , prescribed by Ma et al. (2016a), using the FIRE (Feedback in Realistic Environment) cosmological zoom-in simulations (Hopkins et al. 2014). This suite of simulations reaches a spatial resolution as high as 1-10 pc, 1-2 orders of magnitude better than that in conventional large-volume cosmological hydrodynamic simulations (e.g. Davé et al. 2013), and enable a more realistic and self-consistent treatment of multiphase ISM, star formation, galactic winds and stellar feedback, than those in the semi-analytic models (see Somerville & Davé 2015, for a recent review). As a result, these zoom-in simulations are capable of reproducing many observed relations, e.g., the galaxy stellar mass functions, the evolution of cosmic SFR density and specific SFR (sSFR). The consistency between the shapes of our MZR and that by Ma et al. (2016a) indicates that small-scale physical processes are important in producing the cumulative effects of galactic feedback, and hence high resolution simulations are needed for accurate results.

Aiming at testing the validity of the FMR using the M08-based sample at  $z \sim 1.84$  compiled in this work, we calculate the predicted values for metallicity from the measurements of  $M_*$  and SFR, according to Eq. (2) in Mannucci et al. (2011), who extended the FMR to masses down to  $10^{8.3} M_\odot$ . Figure 9 shows the difference between the FMR-predicted values and the actual measurements of metallicity. We find that the entire M08-based sample is consistent with the FMR given the measurement uncertainties and intrinsic scatter. The median offset for the entire sample is 0.07 dex, with median uncertainty 0.14 dex. The median offset and uncertainty for our galaxy sample analyzed in this work are  $-0.07$  dex and 0.12 dex, respectively. Among other individual datasets, the only one that shows a marginally significant deviation from the FMR is that by Wuyts et al. (2012), which has median offset 0.22 dex and median uncertainty 0.11 dex. We speculate that it is due to the fact that the Wuyts et al. (2012) dataset is the only one in our compiled M08-based sample which relies solely upon the  $[\text{N II}]/\text{H}\alpha$  flux ratio, which can be a biased tracer of metallicity at high- $z$  (Shapley et al. 2015; Sanders et al. 2015). This result reiterates the necessity of combining multiple EL flux ratio diagnostics simultaneously in order to suppress the systematics associated with local calibrations in the accurate measurement of high- $z$  metallicity.

<sup>9</sup> available at <https://github.com/jmeyers314/linmix>



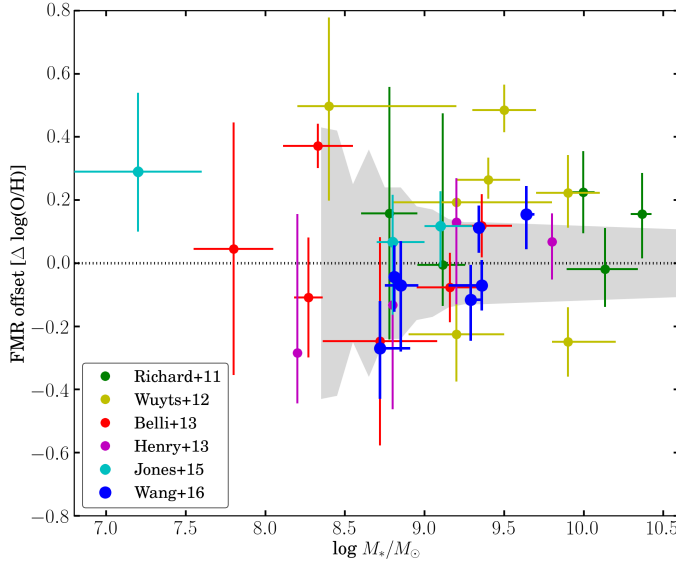


FIG. 9.—: The offset from the fundamental metallicity relation first proposed by Mannucci et al. (2010) for the  $z \sim 1.8$  metallicity measurements derived assuming the Maiolino et al. (2008) calibrations. The color-coding for the points is the same as in Figure 7. The shaded region shows the intrinsic scatter of the FMR extended to low-mass regime given by Mannucci et al. (2011).

TABLE 5  
GAS-PHASE METALLICITY GRADIENTS  
MEASURED BY TWO DIFFERENT METHODS.

ID	Metallicity gradient [dex/kpc]	
	Individual spaxel	Radial annulus
01422	$0.02 \pm 0.08$	$0.06 \pm 0.04$
02389	$-0.07 \pm 0.04$	$-0.13 \pm 0.03$
02806	$-0.01 \pm 0.02$	$0.04 \pm 0.02$
03746	$-0.03 \pm 0.03$	$0.04 \pm 0.02$
04054	$-0.04 \pm 0.02$	$-0.07 \pm 0.02$
05709	$-0.22 \pm 0.05$	$-0.19 \pm 0.06$
05732	$0.06 \pm 0.05$	$0.08 \pm 0.02$
05811	$-0.18 \pm 0.08$	$-0.40 \pm 0.07$
05968	$-0.01 \pm 0.02$	$0.02 \pm 0.01$
07255	$-0.16 \pm 0.03$	$-0.21 \pm 0.03$

In this section, we present and discuss in depth the new and unique results obtained from the spatially resolved analysis of our sample, beyond the reach of the conventional integrated measurements in Section 5. The high spatial resolution maps of gas-phase metallicity from *HST* grism spectroscopy are described in Sect 6.1, and the maps of EL kinematics undertaken with ground-based IFU spectrographs are presented in Section 6.2. We make notes on individual galaxies in Section 6.3.

### 6.1. Gas-phase metallicity maps at sub-kpc resolution

To estimate the spatial distribution of the constrained parameters we apply the analysis described in Section 5 to each individual spaxel in the EL maps. In Figure 10, we show the maps of best-fit metallicity and the corresponding conservative uncertainty (the larger side of asymmetric error bars given by the Bayesian inference) measured for galaxies in our metallicity gradient sample. We bin spaxels 2 by 2 to regain the native WFC3/IR pixel scale. In the maps, we only include spaxels with at least one EL (primarily  $H\alpha$  or  $[O\ III]$ ) detected with  $\geq 5\sigma$  significance. The metallicity gradients measured

from these individual spaxels are shown in the right column of Figure 10. We also measure metallicity gradients using another independent method, i.e., via radial binning. The range of each radial annulus is determined by requiring its SNR  $\geq 15$ . In order to avoid biasing toward low metallicities, we put the SNR threshold on  $H\alpha$  whenever available (i.e.  $z \lesssim 1.5$ ) instead of on  $[O\ III]$ , since the latter correlates tightly with  $12 + \log(O/H)$ , in the lower branch of  $[O\ III]/H\beta$ , where the majority of our sample resides. Based upon our SNR threshold criterion, three sources in our sample, IDs 02918, 05422 and 07058, do not have enough spaxels to constitute a trustworthy metallicity map, and therefore were removed from the spatially resolved analysis. So in total, we present 10 star-forming galaxies whose metallicity gradients (in unit of dex/kpc) are measured. Linear fits to the radial gradients are given in Table 5. We note that the metallicity maps are not always symmetric around the center of stellar mass. Therefore, while gradients are a useful statistic to describe the overall behavior and to compare with numerical simulations, they do not contain all the available information about the apparent diversity of morphologies. Thus, it is helpful to keep the 2D maps in mind while interpreting the gradients.

As a result, seven galaxies have gradients consistent with being flat at  $2\sigma$ , among which three show almost uniformly distributed metals (IDs 02806, 03746 and 05968), two have marginally positive gradients (IDs 01422 and 05732), and the other two display mildly negative gradients (IDs 02389 and 04054). Apart from these, the other three galaxies in our sample have very steep negative gradients: IDs 07255, 05709 and 05811. In particular, after the lensing magnification correction, the stellar mass of galaxy ID 05709 is estimated to be  $\sim 10^{7.9} M_\odot$ . This is for the first time sub-kpc scale spatially resolved analysis has been done on such low-mass systems at high- $z$ .

In order to verify that our results are not contaminated by ionizing radiation from AGN, we examine spatially resolved BPT diagrams. This approach is only possible for sources at  $z \lesssim 1.5$ , where  $H\alpha$  is detectable. Furthermore, due to the low spectral resolution of *HST* grisms, we slightly modify the BPT diagram to plot flux ratio of  $[O\ III]/H\beta$  as a function of  $[S\ II]/H\alpha + [N\ II]$ , using the “vanilla” value for  $[N\ II]/H\alpha$ . As Figure 11 shows, we do not identify any strong correlation between source plane galactocentric radius and deviation from the  $HII$  region loci, suggesting that there is no AGNs hidden in the center of these galaxies. Moreover, the integrated EL fluxes of all three galaxies lie in close proximity to the extreme starburst model prescribed by Kewley et al. (2006), which confirms our working assumption that AGN contributions are negligible for our sample, except in one case (i.e. ID 02607, see Figure 6).

#### 6.1.1. Cosmic evolution of metallicity gradients

We plot the evolution of metallicity gradient across cosmic time in Figure 12. Here we include all existing high- $z$  gradient measurements obtained with sufficiently high spatial sampling, i.e. finer than kpc resolution in the source plane. Together with the 10 new metallicity gradient measurements from the individual spaxel method presented here, in Figure 12, we show one highly magnified galaxy of an interacting triple at  $z=1.85$  analyzed in our previous work (Jones et al. 2015b). We also include 11 measurements by Leethochawalit et al. (2015) (5 isolated and 6 merging), 4 by Jones et al. (2013) (3 isolated and 1 merging), 1 sub- $L_*$  post-merger at  $z \sim 1$  by Frye et al. (2012), 1 isolated grand-design spiral at

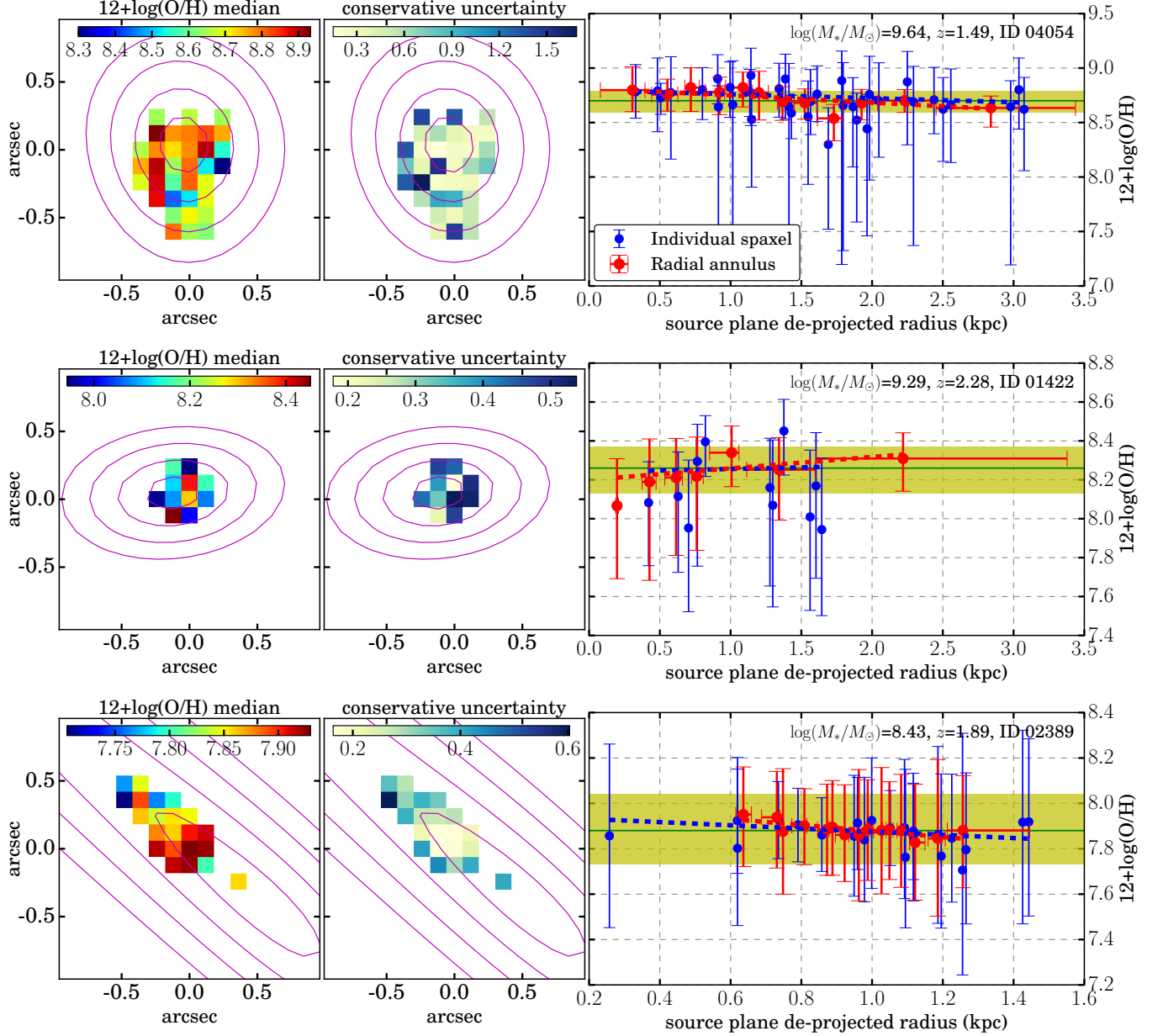


FIG. 10.— Maps of metallicity constraints (median value in the left column and conservative uncertainty — the larger side of asymmetric  $1\text{-}\sigma$  error bars — in the center column) and plots of radial metallicity gradients in the right column. Note that unlike what we show in the combined EL maps (Figures 3 and 4), here the metallicity maps are rebinned to a scale of  $0''.13$ , corresponding to the native resolution of WFC3/IR. In the maps, only spaxels with the signal-to-noise ratio of  $\text{H}\alpha$  or  $[\text{O III}]$  larger than 5 are shown. The same source plane de-projected galactocentric radii that are denoted by black contours in Figures 3 and 4 are represented by magenta contours, with the only difference being that contours are in 1 kpc intervals now. In the panels in the right column, blue and red points correspond to metallicity measurements for individual spaxels and radial annuli, respectively. The radial gradients derived based upon these measurements are shown by the dashed lines in corresponding colors. The yellow band and green horizontal line mark the global constraint on  $12 + \log(\text{O}/\text{H})$ , from integrated line fluxes.

$z=1.49$  by Yuan et al. (2011), and 7 data points from Swinbank et al. (2012) (5 isolated and 2 interacting). The scatter of these high-resolution gradient measurements is  $\sim 0.12$  dex and  $\sim 0.22$  dex at  $z \sim 1.5$  and  $z \sim 2.3$  respectively. The spread of recent gradient measurements from the seeing-limited *KMOS*<sup>3D</sup> survey is also overlaid in Figure 12. Note that the PSF FWHM given by the median seeing of their observations is  $0''.6$ , which results in a 5 kpc resolution at  $z \sim 2$ , twice the size of the half-light radius of an  $L_*$  galaxy at this redshift. We thus only focus on interpreting the sub-kpc res-

olution gradient measurements, in order to avoid possible biases toward null values associated poor spatial sampling (see, e.g., Yuan et al. 2013). Although corrected for beam smearing, the *KMOS*<sup>3D</sup> gradient measurements still display small scatter than that of the high resolution results. When possible, we also divide each data set in terms of sources being isolated or kinematically disturbed (i.e. undergoing merging). In part of our sample, signatures of post-merger tidal remnants can be clearly identified, which strongly indicates that gravitational interaction plays a key role in shaping the spatial distribution

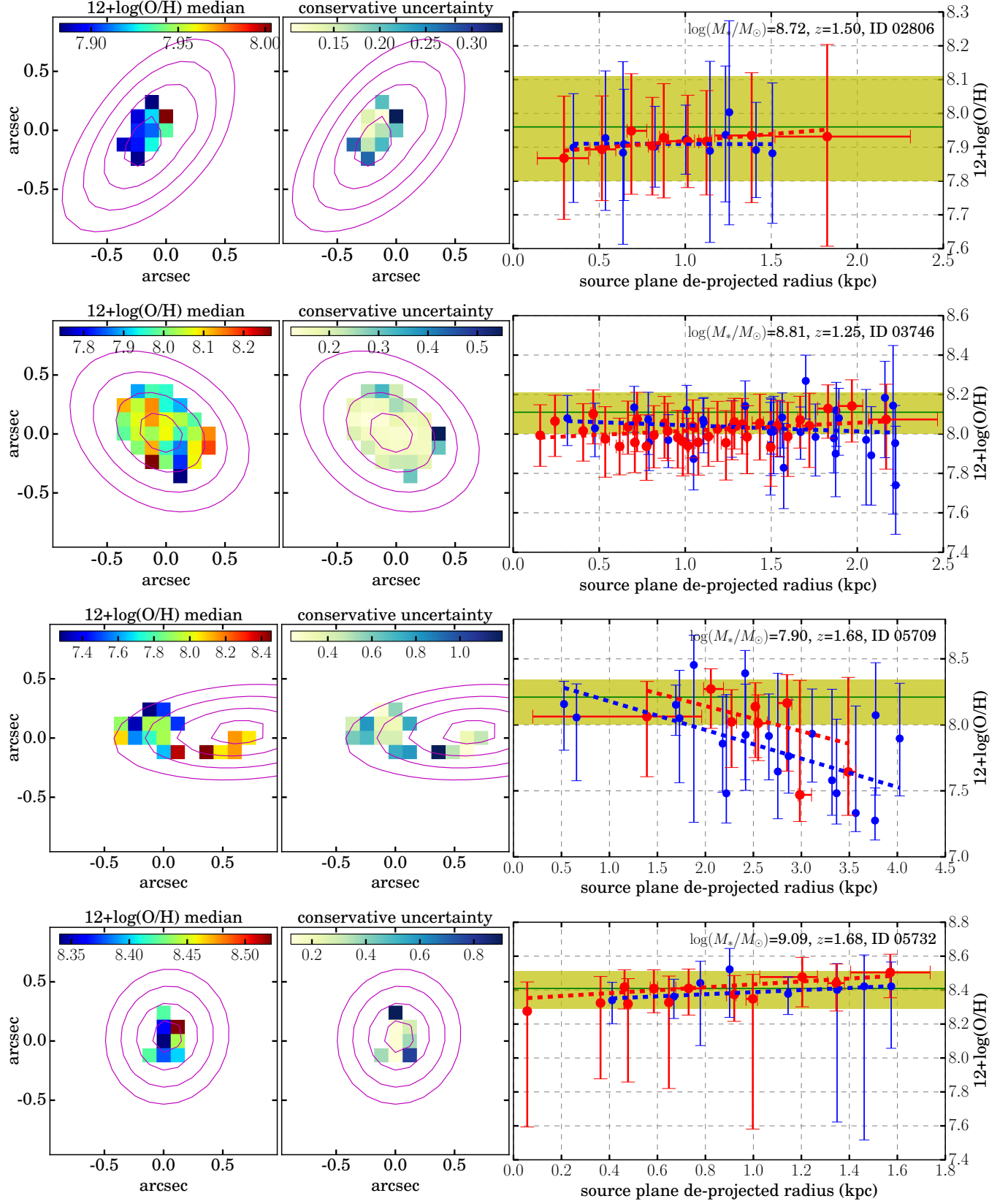


FIG. 10.— (cont.)



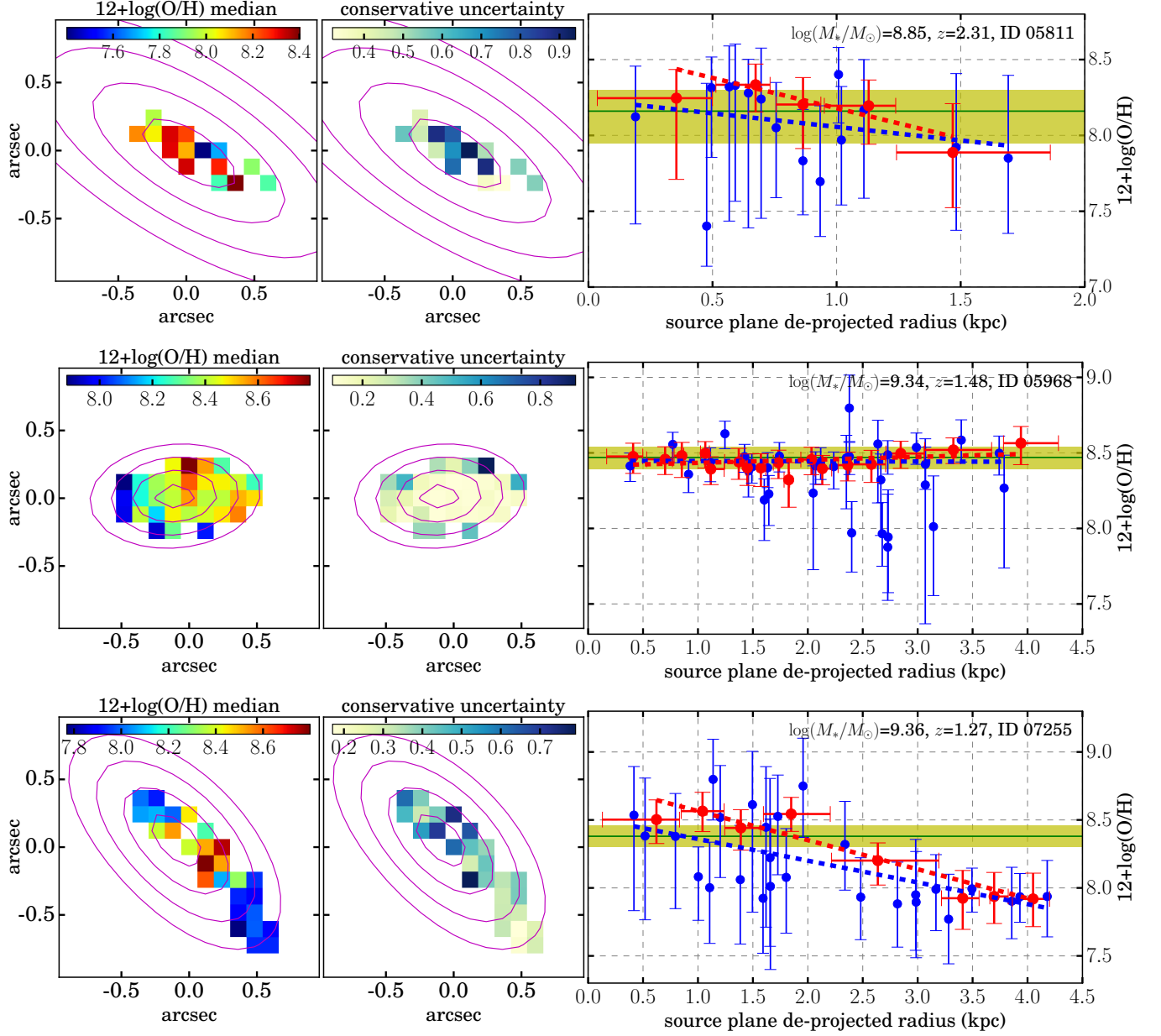


FIG. 10.— (cont.)

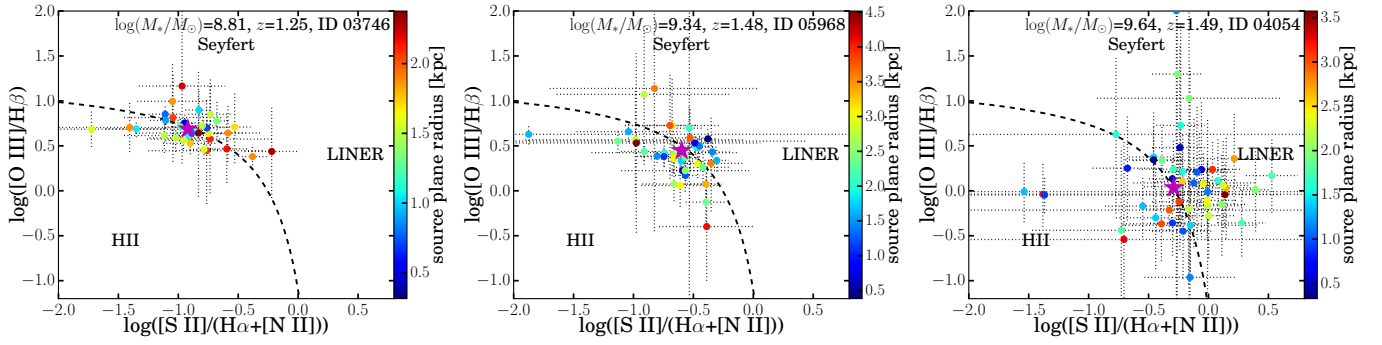


FIG. 11.—: Spatially resolved BPT diagrams for three representative sources in our metallicity gradient sample. The points in each panel correspond to spaxels in the combined EL maps for each source, color coded by the source plane de-projected galactocentric radius. Note that similar to what we show in Figure 10, all spaxels represented by colored points here are rebinned to recover the native WFC3/IR pixel scale ( $0''.13$ ). The magenta star denotes the position where the entire galaxy would lie, calculated from integrated EL fluxes, with the length of the errorbar comparable to the size of the symbol. The dashed curve is adapted from the extreme starburst scenario given by Kewley et al. (2006), assuming the “vanilla” value for  $[N II]/H\alpha$ . Given the  $1-\sigma$  uncertainties, all the spaxels are broadly consistent with being HII regions.

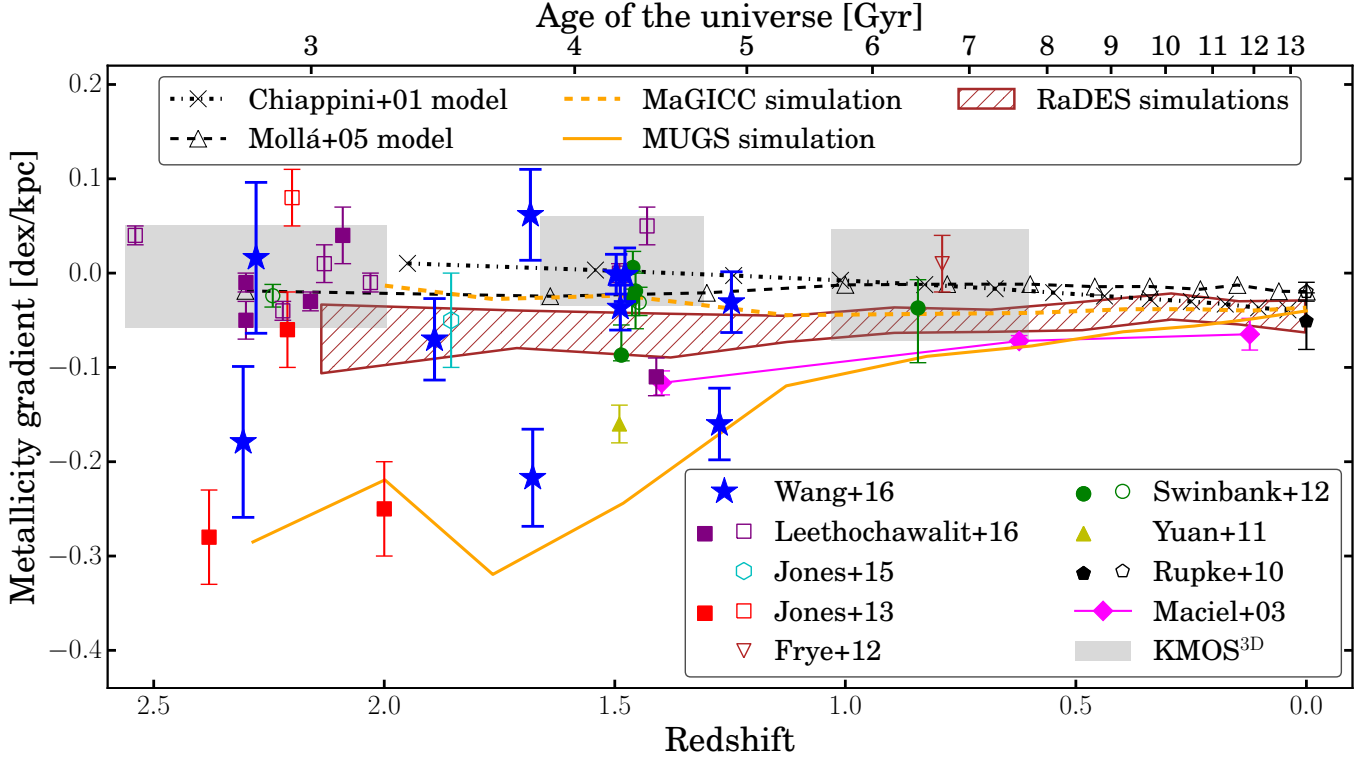


FIG. 12.—: Evolution of metallicity gradients with redshift. The blue stars represent the metallicity gradients measured in this work, and the cyan hexagon is from our previous work (Jones et al. 2015b). We include all the high- $z$  metallicity gradient measurements with sub-kpc resolution in the current literature: lensed galaxies analyzed by Yuan et al. (2011), Frye et al. (2012), Jones et al. (2013), and Leethochawalit et al. (2015), non-lensed galaxies observed with AO (Swinbank et al. 2012). We also show the spread of the  $KMOS^{3D}$  measurements obtained with seeing-limited conditions (Wuyts et al. 2016), the averages of local metallicity gradients (Rupke et al. 2010), and a trend of the Milky Way gradient evolution based upon planetary nebula estimates (Maciel et al. 2003). Except for the blue stars, hollow and filled symbols correspond to interacting/merging and isolated systems, respectively. In addition, the predictions from different analytical chemical evolution models and numerical simulations are also shown as comparisons to the observational results (see Section 6.1 for more details).

of metals in these galaxies (IDs 02806, 03746, and 05968).

In order to compare theoretical predictions with observations, we incorporate into Figure 12 the predictions for the metallicity gradient cosmic evolution from canonical chemical evolution models based upon the “inside-out” disk growth paradigm given by Chiappini et al. (2001, C01) and Mollá & Díaz (2005, M05). C01 constructed a two-phase accretion model for galaxy formation. The first infall period corresponds to the formation of the dark matter halo and the thick disk, with an exponentially declining infall rate at a fixed  $e$ -folding time scale ( $\sim 1$  Gyr). In this phase, the infall of pristine gas is rapid and isotropic, giving birth to a relatively extended profile of star formation. The thin disk is formed in the second phase, where gas is preferentially accreted to the periphery of the disk, with the  $e$ -folding time scale proportional to galactocentric radius (vis-à-vis the “inside-out” growth). M05 proposed their prescription by treating the ISM as a multiphase mixture of hot diffuse gas and cold condensing molecular clouds, and calibrating against different star-formation efficiencies. Here we show the high star-formation efficiency solution. The common Kennicutt-Schmidt law (Kennicutt 1998b; Schmidt 1959) is implemented to estimate star-formation rate. Popular nucleosynthesis prescriptions for type Ia/II SNe and AGB stars are employed by both to construct the yields of dominant isotopes (e.g. H, He, C, N, O, Fe). As shown in Figure 12, the majority of the high- $z$  measurements are generally compatible with both model predictions, with the M05 model showing better consistency, largely ascribed to its more accurate assumptions of ISM structure and more realistic treatment of star formation. Note that these chemical evolution models do not take galactic feedback or radial gas flows into account. So without the consideration of strong physical mechanisms that stir and mix up the enriched gas with relatively unenriched ambient ISM, conventional analytical models are still able to match the metallicity gradients observed at high redshifts.

In addition to the predictions from these analytical models, we also compare our measurements with results from numerical simulations. First, we show the fiducial representative (i.e. the realization g15784) of the McMaster Unbiased Galaxy Simulations (MUGS, Stinson et al. 2010) using the gravitational N-body and SPH code GASOLINE. This simulation employed the “conventional” feedback scheme, i.e. depositing  $\sim 10$ -40% kinetic energy from SN explosions into heating ambient ISM. Furthermore, it adopted a high star-formation threshold (gas particle density  $> 1 \text{ cm}^{-3}$ ), which makes star forming activities highly concentrated in the galaxy center at early stages of disk formation. As a result, a steep gradient (i.e.  $\gtrsim -0.2 \text{ dex/kpc}$ ) is persisting at  $z > 1$  and then significantly flattens with time. Although greatly offset from most of the measurements (and other theoretical predictions) as seen in Figure 12, this particular evolutionary track is actually quite consistent with some of the observed steep gradients (galaxies 07255, 05709, 05811 in our analysis). In comparison, the same galaxy (g15784) is re-simulated with an enhanced feedback prescription of sub-grid physics as part of the Making Galaxies In a Cosmological Context (MaGICC, Stinson et al. 2013) program. A factor of 2.5 difference in the energy output rate from SN feedback diverges the evolutionary track of metallicity gradients given by the MUGS and MaGICC simulations, making the evolution of the MaGICC gradient less dramatic and more similar to the M05 model solution. This demonstrates that amplifying the feedback strength can significantly flatten the metallicity gradients

in early star-forming galaxies (Gibson et al. 2013). This is actually adopted by some recent studies (Leethochawalit et al. 2015; Wuyts et al. 2016) to explain their metallicity gradient measurements, which are also shown in Figure 12. However, as we have already mentioned, this is not the only explanation possible.

Lastly, in Figure 12, we also show another set of numerical results, from an ensemble of 19 galaxies in different environments (9 in isolated field and 10 in loose groups), simulated by the RAMSES Disk Environment Study (RaDES, Few et al. 2012) using the adaptive mesh refinement code RAMSES. Unlike what is shown for the MUGS and MaGICC simulations (i.e. just one single realization), the full range of the whole suite of RaDES simulations is marked in Figure 12, which provides a statistical comparison sample, although the scatter of the RaDES simulations cannot capture the scatter seen in actual measurements. Generally speaking, the evolutionary trends of the abundance gradients of RaDES galaxies lie somewhat in-between the two tracks of g15784 in MUGS and MaGICC simulations, but still quite consistent with the majority of the high- $z$  gradient measurements. Because no mass loading is assumed in the RaDES simulations, outflows do not play any role in this observed consistency. Moreover, since a lower star-formation threshold is used ( $> 0.1 \text{ cm}^{-3}$ ), which is similar to what the analytical chemical evolution models assumed, star formation in RaDES galaxies occurs more uniformly in the early disk formation phase. This strongly suggests that having more extended star formation can also result in galaxies possessing shallower gradients, confirming the tight link between star formation profile and metallicity gradient (Pilkington et al. 2012).

### 6.1.2. Mass dependence of metallicity gradients

In Figure 13, we plot metallicity gradient as a function of stellar mass for a subset of all gradient measurements shown in Figure 12, where reliable  $M_*$  estimates are available. Notably, in the low-mass regime of  $M_* \lesssim 10^9 M_\odot$ , all the current sub-kpc resolution metallicity gradient measurements at  $z \gtrsim 1$  are obtained from our work, and our measurements even extend to below  $10^8 M_\odot$ . This illustrates the power of combining the space-based grism spectroscopy with gravitational lensing as a means to recover reliable metal abundance distributions in early star-forming disks in the low-mass regime. From these measurements, we can tentatively observe an intriguing correlation between  $M_*$  and metallicity gradient, consistent with the galaxy formation “downsizing” picture (see, e.g., Brinchmann et al. 2004; Fontanot et al. 2009) that more massive galaxies are more evolved and their metal distributions flatter since they are in a later phase of disk mass assembly where star formation occurs in a more coherent mode throughout the disk. This is also seen (although in a much smaller  $M_*$  range) in the RaDES simulations that more massive galaxies tend to have shallower (less negative) gradients (Few et al. 2012).

As a comparison, we include the linear fit to the low spatial resolution metallicity gradient measurements from the *KMOS*<sup>3D</sup> survey in the similar  $z$  range (see Figure 6 in Wuyts et al. 2016). Because the mass completeness limit for the *KMOS*<sup>3D</sup> data set is  $\sim 10^{9.7} M_\odot$ , beyond that limit we show the extrapolation of the linear fit. As we have already seen in Figure 7, our analysis is highly complementary to the *KMOS*<sup>3D</sup> result in terms of  $M_*$  associated with metallicity gradient measurements, albeit at very different spatial resolution. However, the low-mass extension of the *KMOS*<sup>3D</sup> fitting relation systematically deviates from our gradient measurements (espe-



TABLE 6  
VELOCITY DISPERSION FOR OBJECTS WITH GROUND-BASED IFU DATA  
AVAILABLE.

ID	Instrument	Emission feature	$\sigma_{\text{obs}}$ (km s $^{-1}$ ) <sup>a</sup>	$\sigma_{\text{int}}$ (km s $^{-1}$ ) <sup>b</sup>
02389	<i>MUSE</i>	C III]	63 $\pm$ 4	19 $^{+11}_{-19}$
02806	<i>MUSE</i>	[O II]	42 $\pm$ 3	24 $\pm$ 5
03746	<i>MUSE</i>	[O II]	74 $\pm$ 9	63 $^{+10}_{-11}$
04054	<i>MUSE</i>	[O II]	68 $\pm$ 6	59 $\pm$ 7
05709	<i>KMOS</i>	[O III]	41 $\pm$ 7	17 $^{+13}_{-17}$
05968	<i>KMOS</i>	[O III]	85 $\pm$ 6	76 $\pm$ 7

<sup>a</sup> Not corrected for instrument resolution

<sup>b</sup> Corrected intrinsic velocity dispersion

cially the ones below  $10^9 M_{\odot}$ ), and is also in contrast with the “downsizing” picture. This can be attributed to possible biases associated with the low spatial resolution of the *KMOS*<sup>3D</sup> observations and the small mass range of their sample.

We also show the  $2\text{-}\sigma$  spread of the galaxy metallicity gradients from the FIRE cosmological zoom-in simulations (Ma et al. 2016b). This suite of simulations has been used to study the MZR as we discussed in Section 5 (Ma et al. 2016a). The simulations show a diversity of gradient behaviors at four redshift epochs ( $z=0, 0.8, 1.4, 2.0$ ). At  $z = 2$ , the scatter of their simulation results can reach 0.2 dex when radial gradients are measured in the central 2 kpc region, quite consistent with the scatter ( $\sim 0.22$  dex) seen in high-resolution gradient measurements at  $z \sim 2.3$ . However, their simulations still cannot reproduce some of the steep gradient measurements, especially some of our results in the low-mass regime. According to Ma et al. (2016b), galactic feedback from intense starburst episodes on 0.1-1 Gyr time scale can effectively flatten metallicity gradients. Hence for the galaxies measured with steep metallicity gradients, galactic feedback must be of limited influence, at least in affecting metal distribution in HII regions on such short time scale.

In short, we observe a tentative mass dependence of metallicity gradients that low-mass galaxies have steeper gradients compared with high-mass galaxies at high redshifts. A larger set of accurately measured and uniformly analyzed metallicity gradients well sampled in  $M_*$  (in the regime of  $M_* \lesssim 10^9 M_{\odot}$  in particular) is needed in order to fully explore the validity of the metallicity gradient “downsizing” picture and whether galactic feedback plays a crucial role in altering this picture.

### 6.2. Emission line kinematics

Kinematic information is valuable for interpreting metallicity gradients and their evolution. Where available, we derive gas-phase kinematics for galaxies observed with *MUSE* or *KMOS*, by fitting the strongest available emission feature. This is typically C III]  $\lambda\lambda 1907, 1909$ , [O II], or [O III]  $\lambda\lambda 4960, 5008$ . We fit each doublet as a sum of two Gaussian components with equal velocity dispersion ( $\sigma$ ) and redshift, plus a constant continuum level. The fits are weighted by the inverse-variance spectrum estimated from sky regions in the data cubes. Best-fit values for the spatially integrated spectra are given in Table 6. The intrinsic velocity dispersion  $\sigma_{\text{int}}$  is calculated by subtracting the instrument resolution (determined from sky ELs) in quadrature from the observed best-fit velocity dispersion  $\sigma_{\text{obs}}$ . For C III] and [O II] we determine the best-fit doublet ratio of the integrated spectra, while the [O III] doublet ratio is fixed to its intrinsic atomic value. For individual spaxels, we fix the doublet ratios to their integrated values in order to avoid spurious fits. We also spatially smooth

each data cube with a 3-pixel ( $0''.6$ ) kernel to improve SNR in the diffuse extended regions. We attempt to fit every spaxel in the smoothed data cube surrounding each object. Bad fits are rejected on the basis of low SNR and unrealistic values (typically requiring  $20 \text{ km s}^{-1} < \sigma_{\text{obs}} < 200 \text{ km s}^{-1}$ , and velocities within  $200 \text{ km s}^{-1}$  of the integrated systemic value). Figure 14 shows velocity fields of each source from all spaxels with acceptable EL fits. We discuss the kinematics of individual galaxies in the context of their metallicity gradients and other properties in Section 6.3.

### 6.3. Notes on individual galaxies

In this section, we highlight some noteworthy sources in our metallicity gradient sample. For reference, all the measurements of global properties, metallicity gradients, and kinematic decompositions in these sources are presented in Tables 2, 5, and 6, respectively.

**02389:** This is the most highly magnified source in our sample, with  $\mu = 43.25^{+7.03}_{-5.78}$ , given by the SHARON & JOHNSON version 3 model. It is one of two connected double images that straddle the lensing critical curve at source redshift  $z = 1.89$ , forming a highly sheared arc more than 5 arcsec. This is confirmed by the velocity shear of  $\sim 50 \text{ km s}^{-1}$  detected in both components of the fold arc. After lensing correction, the stellar mass of this source is  $\log(M_*/M_{\odot}) = 8.43^{+0.06}_{-0.07}$ , consistent with the very low dynamical mass indicated by both the velocity shear and integrated velocity dispersion. The relatively high  $V/\sigma \gtrsim 2$  of this source shows that it is a disk galaxy. The metallicity gradient is mildly negative, i.e.,  $-0.07 \pm 0.04$ , consistent with the prediction by analytical chemical evolution models for disk galaxies at similar redshifts.

**02806:** This is a star-forming galaxy at  $z = 1.50$  with very flat metallicity gradient of  $-0.01 \pm 0.02$ . We removed the contamination from a marginally overlapping cluster member (with greenish color in the color composite stamp shown in Figure 4). Note that such a process would not be possible with seeing-limited data quality. The *MUSE* observation does not spatially resolve this source well, and its [O II] emission line is detected with low SNR near the red edge of *MUSE* spectral coverage. A tentative velocity shear is seen at low significance. The faint fuzzy elongation to the south-east of this galaxy suggests a likely recent merging event, which accounts for its gradient being flat. (See sources 03746 and 05968 for more outstanding cases where post-merger features can be identified.) The global metallicity of this source ( $12 + \log(\text{O}/\text{H}) = 7.91^{+0.16}_{-0.16}$ ) is slightly lower than the average given by the MZR at similar  $M_*$ , in agreement with the study of Michel-Dansac et al. (2008) suggesting that the infall of low-metallicity gas during mergers reduces the average metallicity.

**03746:** This is an intense starburst galaxy with  $\text{SFR} = 10.95^{+1.61}_{-0.58} M_{\odot}/\text{yr}$  at  $z = 1.25$ , as revealed by the BPT diagram in Figure 11. It also displays a flat metallicity gradient of  $-0.03 \pm 0.03$ . Moreover the EL maps of this source show highly clumpy star-forming regions, which do not have direct correspondence in the broad-band photometry. The *MUSE* observation gives a velocity shear of  $\sim 20 \text{ km s}^{-1}$ , and an integrated dispersion of  $\sim 60 \text{ km s}^{-1}$ , which suggests that this system is dispersion dominated ( $V/\sigma < 1$ ). There is an extended feature southward to the source center, which is probably a tidal remnant, consistent with this source being in a recent post-merger phase. Besides, given the irregularities seen in the central region of the  $M_*$  map, this system is still long be-

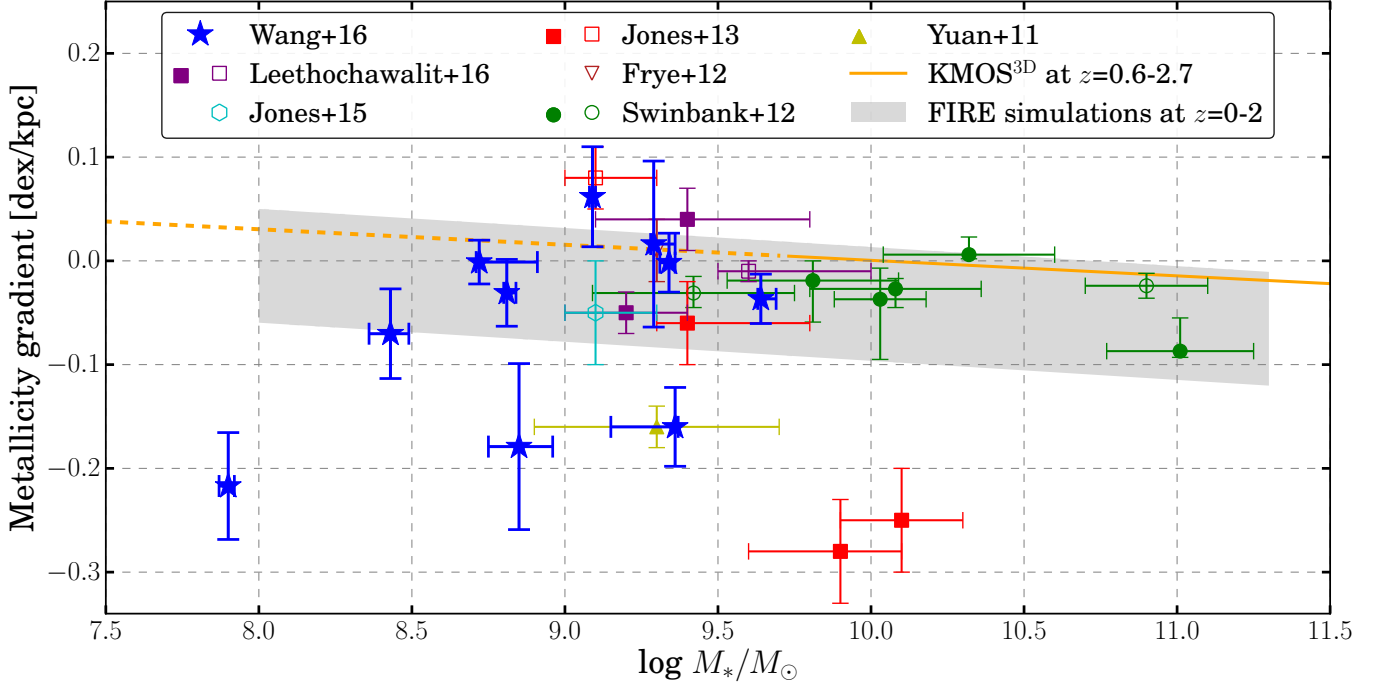


FIG. 13.—: Observed metallicity gradients as a function of  $M_*$ . The color-coding of symbols is the same as in Figure 12. Except for the blue stars, hollow and filled symbols correspond to interacting/merging and isolated systems, respectively. Most of the Leethochawalit et al. (2015) gradient measurements do not have reliable  $M_*$  estimates due to the lack of near IR photometry. The orange line shows a simple linear fit to the metallicity gradient results at  $z=0.6-2.7$  from the *KMOS*<sup>3D</sup> observations (Wuyts et al. 2016). The mass complete range is marked by the solid line, and the dashed line is only an extension from the linear fit. The  $2\text{-}\sigma$  interval of the FIRE simulation results presented by Ma et al. (2016b) is shown as the shaded region. The only work to date that presents metallicity gradient measurements in the  $M_* \lesssim 10^9 M_\odot$  regime is this work, which even extends to below  $10^8 M_\odot$ .

fore dynamical relaxation, and we likely capture this merger system almost right after the coalescence. Therefore we conclude that the flat metallicity gradient recovered in this system is from tidal interactions. See galaxy 05968 below for an even more outstanding post-merger case.

**04054:** This is one of the multiple images of the SN Refsdal host galaxy, a highly magnified almost face-on spiral at  $z = 1.49$ . It is labeled as image 1.3 by Treu et al. (2016)<sup>10</sup>. Our global demographic analysis yields the following measurements:  $\log(M_*/M_\odot) = 9.64^{+0.05}_{-0.01}$ ,  $\text{SFR} = 16.99^{+5.71}_{-2.80} M_\odot/\text{yr}$ ,  $12 + \log(\text{O}/\text{H}) = 8.70^{+0.09}_{-0.11}$ . Compared with the previous analysis of this galaxy (Livermore et al. 2015; Rawle et al. 2016, not necessarily the same image though), our work benefits from a more complete wavelength coverage of imaging data (from  $B_{435}$  to  $H_{160}$  bands) at a greater depth. In particular, we used the most updated lens model, optimized specifically for this spiral galaxy (Treu et al. 2016), giving a reliable estimate of the magnification. Rawle et al. (2016) measured  $\text{SFR} \sim 5 M_\odot/\text{yr}$  on image 1.1 of the SN Refsdal host from *Herschel* IR photometry. However their value is derived assuming a magnification factor of  $\mu = 23.0$ , given by an older lens model for MACS1149.6+2223 (Smith et al. 2009). If the magnification value of  $\mu = 8.14$  given by the latest GLAFIC model is adopted instead, their measurement yields  $\text{SFR} \sim 14$ , comparable to our SFR measurement. As a result, our measurements are fairly consistent with the star-formation main se-

quence, the MZR and the FMR at similar redshifts. Given the high  $M_*$  and  $12 + \log(\text{O}/\text{H})$ , this galaxy is significantly evolved by  $z \sim 1.5$  and with the still relatively high SFR, it will likely turn into a galaxy more massive than our Milky Way at  $z = 0$ . According to the kinematic map derived from the *MUSE* spectroscopy of the [O II] EL doublet, we obtain the velocity shear to be  $\sim 110 \text{ km s}^{-1}$  (see also Figure 8 in Grillo et al. (2016)) and the integrated dispersion  $\sim 60 \text{ km s}^{-1}$ , in good agreement with Livermore et al. (2015) who found  $2v_{2.2} = 118 \pm 6 \text{ km s}^{-1}$ , and local (spatially resolved) velocity dispersion of  $50 \pm 10 \text{ km s}^{-1}$ , based upon adaptive optics data.

The measured metallicity gradient for the SN Refsdal host based upon our analysis of image 1.3 is  $-0.04 \pm 0.02$  from the individual spaxel method, and  $-0.07 \pm 0.02$  from the radial annulus method. This is not as steep as the previous result reported by Yuan et al. (2011), i.e.  $-0.16 \pm 0.02$ , measured from the AO assisted *OSIRIS* IFU observation of image 1.1 of the SN Refsdal host, assuming the Pettini & Pagel (2004) [N II]/H $\alpha$  empirical calibration. After the discovery of SN Refsdal, Kelly et al. (2015c) obtained a 1-hr *MOSFIRE* spectra on the nuclear region and the SN explosion site in image 1.1, and measured  $12 + \log(\text{O}/\text{H}) = 8.6 \pm 0.1$  and  $12 + \log(\text{O}/\text{H}) = 8.3 \pm 0.1$ , respectively, using the same [N II]/H $\alpha$  calibration. The spatial offset between these two measurements in terms of galactocentric radius is  $8.2 \pm 0.5 \text{ kpc}$ , given by the GLAFIC model. As a result, the recent *MOSFIRE* observations indicate a gradient being  $-0.04 \pm 0.02$ , highly consistent with our measurements.

Moreover, we obtain a robust constraint on the nebular dust extinction, i.e.,  $A_V^N = 0.72^{+0.23}_{-0.23}$ . Compared with the stellar dust extinction retrieved from SED fitting, i.e.,  $A_V^S = 1.10^{+0.01}_{-0.01}$ ,

<sup>10</sup> It is located at the minimum of its lensing time delay surface, any intrinsic variabilities in the source are observed first in this image. Unfortunately in the case of SN Refsdal, its image is estimated to appear  $\sim 12$  years ago, long before the deployment of WFC3.

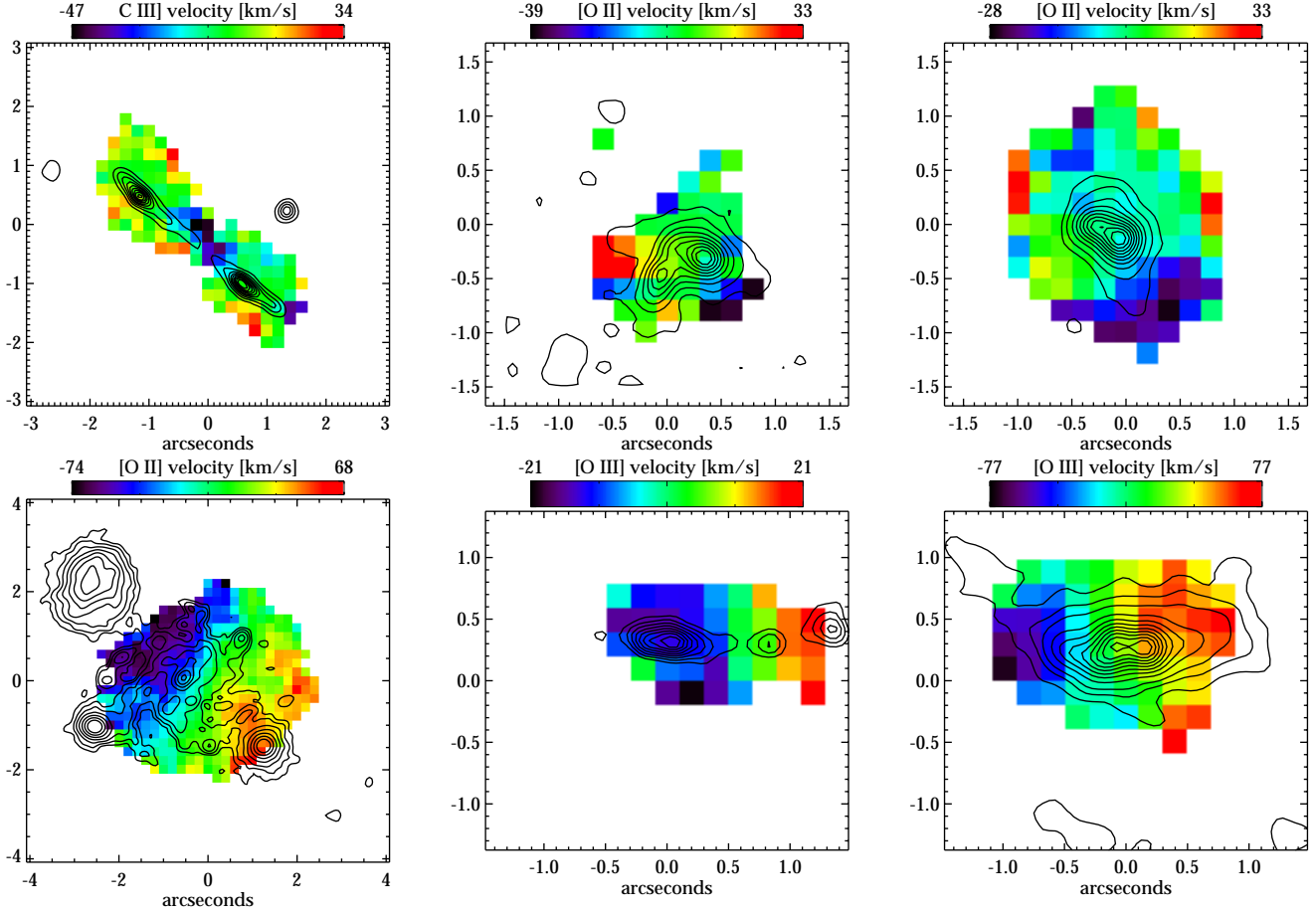


FIG. 14.—: Velocity fields derived from integral field spectra. **Top row:** IDs 02389, 02806, 03746 (left to right) observed with *MUSE*. **Bottom row:** ID 04054 observed with *MUSE* (see also Figure 8 in Grillo et al. (2016)) and IDs 05709, 05968 (left to right) observed with *KMOS*. In all cases, the color scale shows EL velocity centroids in each spaxel, with contours showing *HST/ACS*  $I_{814}$  band continuum. The kinematics map for galaxy 05709 (shown in the lower middle panel) is the first one to date presented for star-forming galaxies with  $M_* \lesssim 10^8 M_\odot$  at  $z \gtrsim 1$ .

we see that in this galaxy, the dust reddening of the ionized gas in HII regions is less severe than that of the stellar continuum (see another source, i.e. ID 05968, from which the same conclusion is drawn). This is different from what has been assumed previously: the color excess of nebular ELs is larger than that of the stellar continuum, i.e.,  $E_S(B - V) = (0.44 \pm 0.03)E_N(B - V)$  (Calzetti et al. 2000). Recently, Reddy et al. (2015) conducted a systematic study of dust reddening using the early observations from the *MOSDEF* survey and found that only 7% of their sample is consistent with  $E_N(B - V) \gtrsim E_S(B - V)/0.44$ . They also discovered that as sSFR (derived from Balmer EL luminosities) increases,  $E_N(B - V)$  diverges more significantly from  $E_S(B - V)$ , with a scatter of  $\sim 0.3$ - $0.4$  dex at  $\text{sSFR} \sim 3 \times 10^{-9} \text{ yr}^{-1}$ , which can account for the difference we see in the total dust attenuation ( $A_V$ ) of the line and continuum emissions of this galaxy.

**05709:** This galaxy has the lowest stellar mass in our metallicity gradient sample ( $\log(M_*/M_\odot) = 7.90^{+0.02}_{-0.03}$ ) after lensing correction<sup>11</sup>. This is for the first time any diffraction-limited metallicity gradient analysis and seeing-limited gas kinematics measurement have been achieved in star-forming galaxies

with such small  $M_*$  at the cosmic noon. This galaxy shows a steep metallicity gradient, i.e.  $-0.22 \pm 0.05$ . The velocity shear is measured to be  $\sim 25 \text{ km s}^{-1}$ , with integrated dispersion being  $\sim 20 \text{ km s}^{-1}$ . The inferred low dynamical mass is consistent with the result from SED fitting mentioned above. In light of the kinematic result  $V/\sigma \gtrsim 1$ , this galaxy is probably a clumpy thick disk. More interestingly, observed from a multi-perspective point of view (see Figure 4), this galaxy shows a clumpy star-forming region at the periphery of its disk (having a distance of 3-4 kpc to the mass center), which has a tremendous amount of line emission, yet very little stellar mass, and has low gas-phase metallicity. Also see galaxy ID 07255 for a more dramatic case.

**05968:** With a tidal tail more appreciably shown to the north-east, and almost perpendicular to the galactic plane, this galaxy (at  $z = 1.48$ ) is safely classified as a post-merger. However note that this identification would not be possible and this source would have been surely classified as a thick disk due to its symmetric morphology and kinematic properties, without the ultra deep *HFF* imaging data. Compared with the  $M_*$  and EL maps of ID 03746, the  $M_*$  map of this galaxy is more smooth and EL maps less clumpy, which indicates that merging has taken place longer before than in source ID 03746. Still the star formation is high, i.e.  $\text{SFR} = 27.95^{+4.70}_{-4.41} M_\odot/\text{yr}$ . The global metallicity ( $12 + \log(\text{O}/\text{H}) = 8.47^{+0.07}_{-0.08}$ )

<sup>11</sup> The magnification value we used in correcting for lensing is  $7.10^{+0.43}_{-0.36}$ , given by the GLAFIC model. The magnification value given by the SHARON & JOHNSON model is similar, i.e.,  $8.43^{+0.73}_{-0.60}$ .



and stellar mass ( $\log(M_*/M_\odot) = 9.34^{+0.02}_{-0.03}$ ) are quite comparable to the MZR at similar redshifts. The metallicity gradient, as expected for systems which have experienced recent mergers (see e.g. Kewley et al. 2010), is constrained to be robustly flat ( $-0.01 \pm 0.02$ ). The EL kinematic analysis yields a velocity shear of  $\sim 120 \text{ km s}^{-1}$ , and an integrated dispersion being  $\sim 75 \text{ km s}^{-1}$ . The velocity field appears typical of intermediate-mass disk galaxies at this redshift, which infers that this post-merger system has been significantly relaxed and regained at least some of the rotation support. This is also one of the few cases where robust constraints on nebular dust extinction can be derived, i.e.  $A_V^N = 0.23^{+0.15}_{-0.13}$ , which is less than the dust extinction of stellar continuum obtained from SED fitting  $A_V^S = 0.60^{+0.20}_{-0.01}$ , as already seen in the case of ID 04054.

**07255:** This galaxy is highly consistent with the SFMS at  $z = 1.27$  formulated by Speagle et al. (2014), with  $\log(M_*/M_\odot) = 9.36^{+0.01}_{-0.21}$  and  $\text{SFR} = 8.73^{+3.53}_{-3.10} M_\odot/\text{yr}$ . By  $z = 0$ , it will turn into a sub-Milky Way-sized galaxy according to Behroozi et al. (2013). Its slightly lower than average metallicity of  $12 + \log(\text{O}/\text{H}) = 8.38^{+0.08}_{-0.08}$  can be a result of low-metallicity gas inflow, following the popular interpretation of the FMR at high  $z$ . This hypothesis would also explain its slightly enhanced SFR. In the metallicity map the inflow can be identified with a bright star-forming clump in the lower right corner, which consists of very little  $M_*$ , but is dominating in the EL maps, similar to the case of galaxy 05709 discussed above. This indicates that in this system, the metal enrichment time scale is much larger than the star-forming time scale, which is in turn much larger than the gas infall time scale. Unfortunately, at the moment we do not have EL kinematic observation on this system.

The metallicity gradient for this galaxy is constrained to be very steep, i.e.  $-0.16 \pm 0.03$  using the individual spaxel method, and  $-0.21 \pm 0.03$  from the radial annulus method. We also dissected this galaxy to separate the star-forming clump from the bulk part and re-do the analysis. It is found that the HII regions closely related to the clump (with at least 2.5 kpc distance to the galaxy mass center) have significantly lower metallicity and provide more than one third of the entire star formation, i.e.,  $12 + \log(\text{O}/\text{H}) = 8.04^{+0.14}_{-0.18}$ ,  $A_V^N < 0.78$ ,  $\text{SFR} = 3.05^{+1.21}_{-1.10}$ , compared with the quantities measured in the bulk of the galaxy (within 2 kpc to the mass center), i.e.,  $12 + \log(\text{O}/\text{H}) = 8.51^{+0.08}_{-0.10}$ ,  $A_V^N = 0.63^{+0.59}_{-0.41}$ ,  $\text{SFR} = 5.78^{+3.41}_{-2.47}$ . In both regions, EL H $\alpha$  is detected at a significance of  $\gtrsim 30\sigma$ . This again demonstrates that measurements of global quantities can be to some extent misleading, and the importance of detailed spatially resolved information. Cases like this galaxy should be exceedingly interesting, since their steep metallicity gradients at such high  $z$  are not predicted by current mainstream numerical simulations (equipped with strong galactic feedback) or conventional analytical chemical evolution models. Through studying systems like this, we can also learn how gas is accreted into the inner disk, and more importantly whether the accretion happens alongside star formation or long before the gas clouds can collapse.

## 7. SUMMARY AND DISCUSSION

### 7.1. Summary

In this paper, we took advantage of the deep *HST* grism spectroscopic data, acquired by the GLASS and SN Refsdal follow-up programs, and obtain gas-phase metallicity maps in high- $z$  star-forming galaxies. The *HST* grisms provide an

un-interrupted window for observing the cosmic evolution of metallicity gradients through  $z$  from 2.3 to 1.2 (corresponding to the age of the universe from roughly 2.5 to 5.5 Gyr), including the redshift range  $1.7 < z < 2.0$ , unattainable from ground-based observations due to low atmospheric transmission. Another advantage of the *HST* grisms is that multiple ELs are available within a single setup and all the potential issues related to combining data with different setups and atmospheric conditions are eliminated. Our main conclusions can be summarized as follows:

1. We presented 10 maps of gas-phase metallicity (as shown in Figure 10), measured at sub-kpc resolution in star-forming galaxies in the redshift range of  $1.2 \lesssim z \lesssim 2.3$ , in the prime field of MACS1149.6+2223. This field is so far one of the deepest fields exposed by *HST* spectroscopy, at a depth of 34(12) orbits of G141(G102), reaching a  $1\text{-}\sigma$  flux limit of  $1.2(3.5) \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ , without lensing correction.
2. We developed a novel Bayesian statistical approach to jointly constrain gas-phase metallicity, nebular dust extinction, and H $\alpha$ -based dust-corrected star formation rate. In determining these quantities, our method does not rely on any assumptions about star-formation history, nor the conversion of dust reddening from stellar phase to nebular phase. Unlike the majority of previous work on deriving metallicity from strong EL calibrations, our method does not compare directly the observed EL flux ratios with the calibrated ones. Instead we work directly with EL fluxes, which effectively avoids using redundant information and circumvents possible biases in the low SNR regime.
3. The metallicity maps we obtained show a large diversity of morphologies. Some maps display disk-like shapes and have metallicity gradients consistent with those predicted by analytical chemical evolution models (i.e. IDs 04054, 01422, 02389, 05732), whereas other disk-shaped galaxies have exceedingly steep gradients (i.e. IDs 07255, 05709, 05811) and bright star-forming clumps at the periphery of their disks. These large-offset star-forming clumps, containing very little stellar mass, are estimated to have lower metallicity than the corresponding global values, which can be interpreted as an evidence for the infall of low metallicity gas enhancing star formation. We also recovered three systems with nearly uniform spatial distribution of metals (i.e. IDs 02806, 03746, 05968). In these systems, we can identify possible signatures of post-merger tidal remnants, which suggests that the observed flat gradients are likely caused by gravitational interactions.
4. Collecting all existing sub-kpc resolution metallicity gradient measurements at high- $z$ , we study the cosmic evolution and mass dependence of metallicity gradients.
  - We found that predictions given by analytical chemical evolution models assuming a relatively extended star-formation profile in the early phase of disk growth can reproduce the majority of observed metallicity gradients at high- $z$ , without involving strong galactic feedback or radial out-

flows. This confirms the tight link between star formation and metal enrichment.

- We tentatively observed an correlation between stellar mass and metallicity gradient, which is in accord with the “downsizing” galaxy formation scenario that more massive galaxies are more evolved into a later phase of disk growth, where they experience more coherent mass assembly and metal enrichment, than lower mass galaxies.

A larger sample of uniformly analyzed metallicity maps well sampled in the full  $M_*$  range (especially in the low mass regime, i.e.,  $M_* \lesssim 10^9 M_\odot$ ) at high redshifts is needed to further investigate the cosmic evolution and mass dependence of metallicity gradients.

5. We presented 6 maps of EL kinematics obtained from ground-based IFU instruments. Thanks to the powerful synergy of lensing magnification and diffraction-limited data, we push for the first time the mass limit of these spatially resolved analysis (the measurements of metallicity gradients and EL kinematics) to below  $10^8 M_\odot$ .
6. We also compiled a sample of 38 global metallicity measurements all derived from the Maiolino et al. (2008) EL calibrations in the current literature (including 13 galaxies presented in this work). This sample, at  $z \sim 1.8$ , spans three orders of magnitude in  $M_*$ . We measured the MZR and tested the FMR using this sample.
  - The slope of the MZR constrained by this sample rules out the momentum-driven wind model by Davé et al. (2012) at  $3-\sigma$  confidence level. Our MZR is consistent with the theoretical prediction given by recent cosmological zoom-in simulations, suggesting that high spatial resolution simulations are favorable in reproducing the metal enrichment history in star-forming galaxies.
  - Given the intrinsic scatter and measurement uncertainties, we do not see any significant offset from the FMR except the subsample of Wuyts et al. (2012). This discrepancy is likely due to the fact that the Wuyts et al. (2012) dataset is the only subsample which relies exclusively on  $[\text{N II}]/\text{H}\alpha$  to estimate metallicity. We therefore advocate to avoid nitrogen lines when estimating gas-phase metallicity in high- $z$  HII regions.

### 7.2. Interpretation of flat/steep metallicity gradients at high- $z$

We now turn to the interpretation of our observed metallicity gradients and maps. In general, there are four physical mechanisms that could flatten a galaxy’s *pre-existing* negative metallicity gradient (not exclusively in high- $z$  scenarios):

- Efficient radial mixing by tidal effects from gravitational interactions, as seen in galaxy IDs 02806, 03746, and 05968 in our sample and the triply imaged arc 4 in our previous work (Jones et al. 2015b).
- Rapid recycling of metal-enriched material by enhanced galactic feedback, as seen in the results of the MaGICC numerical simulations for example.
- Turbulence mixing driven by thermal instability as elucidated by Yang & Krumholz (2012).

- Funnelling of cold streams of low metallicity gas directly into the galaxy center, as suggested by Dekel et al. (2009)<sup>12</sup>

Besides *pre-existing* negative gradients being flattened by these processes, galaxies can also inherit a broadly flat metallicity gradient from relatively extended star-formation profile in the early disk-formation phase and coherent mass assembly for the disk growth. Because there is strong evidence that the mass assembly for Milky Way progenitors takes place uniformly at all radii, maintaining almost the same  $M_*$  profile from  $z \sim 2.5$  to  $z \sim 1$  (van Dokkum et al. 2013; Morishita et al. 2015). However the uncertainties associated with current measurements are still insufficient to distinguish numerical simulations with enhanced feedback prescriptions from conventional analytical chemical evolution models assuming more uniform star formation (see Figure 12).

One perhaps obvious but nevertheless important caveat common to most galaxy evolution work is that we only have access to *cross-sectional* data, i.e., observations of different objects at one specific epoch. Therefore any interpretations of plots like Figure 12, which try to tie these data with *longitudinal* theories (describing the evolution of one specific object), should be made with caution.

Perhaps an even more intriguing interpretation of Figure 12, which is robust against the aforementioned caveat, is that there exist some really steep metallicity gradients at high  $z$ , such as our galaxy IDs 07255, 05709, and 05811. Based upon current measurements, the occurrence rate of high- $z$  metallicity gradients more negative than  $-0.1$  dex/kpc is 6/35. We can derive the following constraints on the physical processes that can contribute to these steep metallicity gradients observed beyond the local universe:

- galactic feedback (including supernova explosions and stellar winds) must be of limited influence (via being confined for instance),
- star-formation efficiency is low, or at least lower than their local analogs (in simulations, this can be achieved by having a higher star-formation threshold or less concentrated molecular gas, as shown by the MUGS runs),
- star formation at early times is highly centrally concentrated, and the mass assembly of the inner regions is much faster than that of the outer regions,
- alternatively, steep gradients are the result of low-metallicity gas falling onto the outskirts of their disks, implying that the dynamic time scale is much shorter than the star-forming time scale, which is much shorter than the metal enrichment time scale, as seen in our galaxy 07255.

We need larger samples with robust measurements in order to more accurately quantify the occurrence rate of these steep metallicity gradients in high- $z$  star-forming disks. The existence of a distinct population of galaxies with steep metallicity gradients, if confirmed by future observations constraining their abundance, is a powerful test of galaxy formation models, because this is not predicted by most current theoretical models. Hence, these galaxies with steep metallicity gradients should be the prime targets for new numerical predictions: future *JWST* observations will constrain their properties in great detail, allowing discriminating tests of current theories.

Furthermore, most current numerical simulations are aimed at reproducing present-day Milky Way analogs, surrounded

<sup>12</sup> This mechanism is invoked to explain the inverted gradients (lower abundances in the inner disk regions) seen in some of the AMAZE/LSD observations (Cresci et al. 2010). But we note in their observations, the near-IR integral field spectrometer *SINFONI* was used in seeing-limited mode.

by a dark matter halo with Virial mass  $\sim 10^{12} M_{\odot}$ . It would be useful to have large samples of simulated sub- $L_*$  galaxies ( $M_*$  below  $5 \times 10^9 M_{\odot}$  at  $z \sim 2$ ) in order to investigate how metals are recycled in these less massive systems and provide a more suitable comparison to our metallicity gradient measurements at high  $z$ . With improvements on both the observational and theoretical sides, we will be able to answer why in these high- $z$  star-forming disk galaxies, the metallicity gradients can be established so early and why they are so steep compared with those found in local analogs.

We acknowledge support by NASA through grant HST-GO-13459 (PI: Treu). This paper is based on observations made with the NASA/ESA Hubble Space Telescope, and utilizes gravitational lensing models produced by PIs Bradač, Natarajan & Kneib (CATS), Merten & Zitrin, Sharon, and Williams, and the GLAFIC and Diego groups. This lens modeling was partially funded by the HST Frontier Fields program conducted by STScI. STScI is operated by the Asso-

ciation of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. The lens models were obtained from the Mikulski Archive for Space Telescopes (MAST). Ground based data have been obtained with the MUSE and KMOS spectrographs at the ESO VLT Telescopes. TJ acknowledges support provided by NASA through Program # HST-HF2-51359 through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. XW is greatly indebted to Ryan Sanders for a careful reading of an early version of this manuscript, and to Sirio Belli, Sandy Faber, Cheng Li, Hui Li, Zhiyuan Li, Shude Mao, Houjun Mo, Alice Shapley, Yiping Shu, Alessandro Sonnenfeld, Lixin Wang, Tao Wang, Anita Zanella, and Zheng Zheng for helpful discussions. XW also acknowledges the warm hospitality at the Tsinghua Center for Astrophysics where a large portion of this paper was written. Last but not the least, XW thanks Dr. Xiao-Lei Meng for her everlasting love and support.

## REFERENCES

- Andrews, B. H., & Martini, P. 2013, *The Astrophysical Journal*, 765, 140
- Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, *Monthly Notices of the Royal Astronomical Society*, 388, 945
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *Publications of the Astronomical Society of the Pacific*, 93, 5
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *The Astrophysical Journal*, 770, 57
- Belli, S., Jones, T. A., Ellis, R. S., & Richard, J. 2013, *The Astrophysical Journal*, 772, 141
- Bianco, F. B., Modjaz, M., Oh, S. M., et al. 2015, eprint arXiv:1505.06213, 54
- Blanc, G. A., Kewley, L. J., Vogt, F. P. A., & Dopita, M. A. 2015, *The Astrophysical Journal*, 798, 99
- Brammer, G. B., Pirzkal, N., McCullough, P., & MacKenty, J. 2014, *Space Telescope WFC Instrument Science Report*
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *The Astrophysical Journal Supplement Series*, 200, 13
- Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *Monthly Notices of the Royal Astronomical Society*, 351, 1151
- Bruzual, G., & Charlot, S. 2003, *Monthly Notices of the Royal Astronomical Society*, 344, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *The Astrophysical Journal*, 533, 682
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *The Astronomical Journal*, 345, 245
- Chabrier, G. 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763
- Chiappini, C., Matteucci, F., & Romano, D. 2001, *The Astrophysical Journal*, 554, 1044
- Coil, A. L., Aird, J., Reddy, N. A., et al. 2015, *The Astrophysical Journal*, 801, 35
- Coleman, M. G., & de Jong, J. T. A. 2008, *The Astrophysical Journal*, 685, 933
- Cowie, L. L., Barger, A. J., & Songaila, A. 2016, *The Astrophysical Journal*, 817, 57
- Cresci, G., Mannucci, F., Maiolino, R., et al. 2010, *Nature*, 467, 811
- Davé, R., Finlator, K., & Oppenheimer, B. D. 2011a, *Monthly Notices of the Royal Astronomical Society*, 416, 1354
- . 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 98
- Davé, R., Katz, N., Oppenheimer, B. D., Kollmeier, J. A., & Weinberg, D. H. 2013, *Monthly Notices of the Royal Astronomical Society*, 434, 2645
- Davé, R., Oppenheimer, B. D., & Finlator, K. 2011b, *Monthly Notices of the Royal Astronomical Society*, 415, 11
- Dekel, A., Birnboim, Y., Engel, G., et al. 2009, *Nature*, 457, 451
- Dopita, M. A., Sutherland, R. S., Nicholls, D. C., Kewley, L. J., & Vogt, F. P. A. 2013, *The Astrophysical Journal Supplement Series*, 208, 10
- Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, *The Astrophysical Journal*, 644, 813
- Few, C. G., Gibson, B. K., Courty, S., et al. 2012, *Astronomy and Astrophysics*, 547, A63
- Finlator, K., & Davé, R. 2008, *Monthly Notices of the Royal Astronomical Society*, 385, 2181
- Fontanot, F., De Lucia, G., Monaco, P., Somerville, R. S., & Santini, P. 2009, *Monthly Notices of the Royal Astronomical Society*, 397, 1776
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *Publications of the Astronomical Society of the Pacific*, 125, 306
- Frye, B. L., Hurley, M., Bowen, D. V., et al. 2012, *The Astrophysical Journal*, 754, 17
- Gibson, B. K., Pilkington, K., Brook, C. B., Stinson, G. S., & Bailin, J. 2013, *Astronomy and Astrophysics*, 554, A47
- Grillo, C., Karman, W., Suyu, S. H., et al. 2016, *The Astrophysical Journal*, 822, 78
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *The Astrophysical Journal Supplement Series*, 197, 35
- Guo, Y., Koo, D. C., Lu, Y., et al. 2016, eprint arXiv:1603.04863, 1603.04863
- Henry, A., Scarlata, C., Domínguez, A., et al. 2013, *The Astrophysical Journal*, 776, L27
- Hopkins, P. F., Kere, D., Onorbe, J., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 581
- Hummer, D. G., & Storey, P. J. 1987, *Monthly Notices of the Royal Astronomical Society*, 224, 801
- Hunt, L., Dayal, P., Magrini, L., & Ferrara, A. 2016a, eprint arXiv:1608.05417, 1608.05417
- . 2016b, eprint arXiv:1608.05418, 1608.05418
- Johnson, T. L., Sharon, K., Bayliss, M. B., et al. 2014, *The Astrophysical Journal*, 797, 48
- Jones, T. A., Ellis, R. S., Jullo, E., & Richard, J. 2010, *The Astrophysical Journal Letters*, 725, L176
- Jones, T. A., Ellis, R. S., Richard, J., & Jullo, E. 2013, *The Astrophysical Journal*, 765, 48
- Jones, T. A., Martin, C., & Cooper, M. C. 2015a, *The Astrophysical Journal*, 813, 126
- Jones, T. A., Wang, X., Schmidt, K. B., et al. 2015b, *The Astronomical Journal*, 149, 107
- Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, *The Astrophysical Journal*, 736, 104
- Juneau, S., Bouchaud, F., Charlot, S., et al. 2014, *The Astrophysical Journal*, 788, 88
- Karman, W., Grillo, C., Balestra, I., et al. 2016, *Astronomy and Astrophysics*, 585, A27
- Kawamata, R., Oguri, M., Ishigaki, M., Shimasaku, K., & Ouchi, M. 2015, arXiv.org, arXiv:1510.06400
- Kelly, B. C. 2007, *The Astrophysical Journal*, 665, 1489
- Kelly, P. L., Rodney, S. A., Treu, T., et al. 2015a, arXiv.org, arXiv:1512.04654
- . 2015b, *Science*, 347, 1123
- Kelly, P. L., Brammer, G. B., Selsing, J., et al. 2015c, arXiv.org, arXiv:1512.09093

- Kennicutt, R. C. J. 1998a, *Annual Review of Astronomy and Astrophysics*, 36, 189
- . 1998b, *The Astrophysical Journal*, 498, 541
- Kewley, L. J., & Ellison, S. L. 2008, *The Astrophysical Journal*, 681, 1183
- Kewley, L. J., Groves, B., Kauffmann, G., & Heckman, T. M. 2006, *Monthly Notices of the Royal Astronomical Society*, 372, 961
- Kewley, L. J., Rupke, D. S. N., Zahid, H. J., Geller, M. J., & Barton, E. J. 2010, *The Astrophysical Journal*, 721, L48
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *The Astrophysical Journal Supplement Series*, 197, 36
- Kriek, M. T., van Dokkum, P. G., Labbé, I., et al. 2009, *The Astrophysical Journal*, 700, 221
- Kümmel, M., Walsh, J. R., Pirzkal, N., Kuntschner, H., & Pasquali, A. 2009, *Publications of the Astronomical Society of the Pacific*, 121, 59
- Leethochawalit, N., Jones, T. A., Ellis, R. S., et al. 2015, *arXiv.org*, [arXiv:1509.01279](https://arxiv.org/abs/1509.01279)
- Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y.-j. 2013, *The Astrophysical Journal*, 772, 119
- Livermore, R. C., Jones, T. A., Richard, J., et al. 2015, *Monthly Notices of the Royal Astronomical Society*, 450, 1812
- Lotz, J. M., Koekemoer, A. M., Coe, D., et al. 2016, 1605.06567
- Lu, Y., Blanc, G. A., & Benson, A. 2015a, *The Astrophysical Journal*, 808, 129
- Lu, Z., Mo, H. J., & Lu, Y. 2015b, *Monthly Notices of the Royal Astronomical Society*, 450, 606
- Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016a, *Monthly Notices of the Royal Astronomical Society*, 456, 2140
- Ma, X., Hopkins, P. F., Feldmann, R., et al. 2016b, *eprint arXiv:1610.03498*, 1610.03498
- Maciel, W. J., Costa, R. D. D., & Uchida, M. M. M. 2003, *Astronomy and Astrophysics*, 397, 667
- Madau, P., & Dickinson, M. 2014, *Annual Review of Astronomy and Astrophysics*, 52, 415
- Maiolino, R., Nagao, T., Grazian, A., et al. 2008, *Astronomy and Astrophysics*, 488, 463
- Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, *Monthly Notices of the Royal Astronomical Society*, 408, 2115
- Mannucci, F., Salvaterra, R., & Campisi, M. A. 2011, *Monthly Notices of the Royal Astronomical Society*, 414, 1263
- Martin, C. L., Shapley, A. E., Coil, A. L., et al. 2012, *The Astrophysical Journal*, 760, 127
- Mason, C. A., Treu, T., Fontana, A., et al. 2016, *eprint arXiv:1610.03075*, 1610.03075
- Masters, D., Faisst, A., & Capak, P. 2016, *eprint arXiv:1605.04314*, 1605.04314
- Meneghetti, M., Natarajan, P., Coe, D., et al. 2016, *eprint arXiv:1606.04548*, 1606.04548
- Michel-Dansac, L., Lambas, D. G., Alonso, M. S., & Tissera, P. 2008, *Monthly Notices of the Royal Astronomical Society*, 386, L82
- Mollá, M., & Díaz, A. I. 2005, *Monthly Notices of the Royal Astronomical Society*, 358, 521
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2015, *eprint arXiv:1510.02106*, 1510.02106
- Morishita, T., Ichikawa, T., Noguchi, M., et al. 2015, *The Astrophysical Journal*, 805, 34
- Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 433, 1185
- Moustakas, J., Kennicutt, R. C. J., Tremonti, C. A., et al. 2010, *The Astrophysical Journal Supplement Series*, 190, 233
- Oguri, M. 2010, *Publications of the Astronomical Society of Japan*, 62, 1017
- Oke, J. B., & Gunn, J. E. 1983, *Astrophysical Journal*, 266, 713
- Pérez-Montero, E. 2014, *Monthly Notices of the Royal Astronomical Society*, 441, 2663
- Pettini, M., & Pagel, B. E. J. 2004, *Monthly Notices of the Royal Astronomical Society*, 348, L59
- Pilkington, K., Few, C. G., Gibson, B. K., et al. 2012, *Astronomy and Astrophysics*, 540, A56
- Pilyugin, L. S., Vilchez, J. M., & Thuan, T. X. 2010, *The Astrophysical Journal*, 720, 1738
- Queyrel, J., Contini, T., Kissler-Patig, M., et al. 2012, *Astronomy and Astrophysics*, 539, A93
- Rawle, T. D., Altieri, B., Egami, E., et al. 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 1626
- Reddy, N. A., Kriek, M. T., Shapley, A. E., et al. 2015, *The Astrophysical Journal*, 806, 259
- Richard, J., Jones, T. A., Ellis, R. S., et al. 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 643
- Rupke, D. S. N., Kewley, L. J., & Chien, L. H. 2010, *The Astrophysical Journal*, 723, 1255
- Sánchez, S. F., Rosales-Ortega, F. F., Iglesias-Páramo, J., et al. 2014, *Astronomy and Astrophysics*, 563, A49
- Sanders, R. L., Shapley, A. E., Kriek, M. T., et al. 2015, *The Astrophysical Journal*, 799, 138
- . 2016, *eprint arXiv:1606.04107*, 1606.04107
- Schmidt, K. B., Treu, T., Brammer, G. B., et al. 2014, *The Astrophysical Journal Letters*, 782, L36
- Schmidt, K. B., Treu, T., Bradac, M., et al. 2016, *The Astrophysical Journal*, 818, 38
- Schmidt, M. 1959, *The Astrophysical Journal*, 129, 243
- Shapley, A. E., Reddy, N. A., Kriek, M. T., et al. 2015, *The Astrophysical Journal*, 801, 88
- Sharon, K., & Johnson, T. L. 2015, *The Astrophysical Journal*, 800, L26
- Smith, G. P., Ebeling, H., Limousin, M., et al. 2009, *The Astrophysical Journal*, 707, L163
- Somerville, R. S., & Davé, R. 2015, *Annual Review of Astronomy and Astrophysics*, 53, 51
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *The Astrophysical Journal Supplement Series*, 214, 15
- Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *The Astrophysical Journal*, 795, 165
- Stinson, G. S., Bailin, J., Couchman, H., et al. 2010, *Monthly Notices of the Royal Astronomical Society*, 408, 812
- Stinson, G. S., Brook, C. B., Macciò, A. V., et al. 2013, *Monthly Notices of the Royal Astronomical Society*, 428, 129
- Storey, P. J., & Zeppen, C. J. 2000, *Monthly Notices of the Royal Astronomical Society*, 312, 813
- Stott, J. P., Sobral, D., Swinbank, A. M., et al. 2014, *Monthly Notices of the Royal Astronomical Society*, 443, 2695
- Swinbank, A. M., Sobral, D., Smail, I., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 426, 935
- Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, *The Astrophysical Journal*, 613, 898
- Treu, T., Schmidt, K. B., Brammer, G. B., et al. 2015, *The Astrophysical Journal*, 812, 114
- Treu, T., Brammer, G. B., Diego, J. M., et al. 2016, *The Astrophysical Journal*, 817, 60
- Troncoso, P., Maiolino, R., Sommariva, V., et al. 2014, *Astronomy and Astrophysics*, 563, A58
- van Dokkum, P. G., Leja, J., Nelson, E., et al. 2013, *The Astrophysical Journal*, 771, L35
- Vulcani, B., Treu, T., Schmidt, K. B., et al. 2015, *The Astrophysical Journal*, 814, 161
- Vulcani, B., Treu, T., Nipoti, C., et al. 2016, *eprint arXiv:1610.04615*, 1610.04615
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *The Astrophysical Journal*, 795, 104
- Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, *The Astrophysical Journal*, 799, 209
- Wuyts, E., Rigby, J. R., Sharon, K., & Gladders, M. D. 2012, *The Astrophysical Journal*, 755, 73
- Wuyts, E., Kurk, J., Förster Schreiber, N. M., et al. 2014, *The Astrophysical Journal*, 789, L40
- Wuyts, E., Wisnioski, E., Fossati, M., et al. 2016, *eprint arXiv:1603.01139*, 1603.01139
- Yang, C.-C., & Krumholz, M. R. 2012, *The Astrophysical Journal*, 758, 48
- Yuan, T., Kewley, L. J., & Rich, J. A. 2013, *The Astrophysical Journal*, 767, 106
- Yuan, T., Kewley, L. J., Swinbank, A. M., Richard, J., & Livermore, R. C. 2011, *The Astrophysical Journal*, 732, L14
- Zahid, H. J., Bresolin, F., Kewley, L. J., Coil, A. L., & Davé, R. 2012, *The Astrophysical Journal*, 750, 120
- Zahid, H. J., Dima, G. I., Kudritzki, R.-P., et al. 2014, *The Astrophysical Journal*, 791, 130
- Zahid, H. J., Kewley, L. J., & Bresolin, F. 2011, *The Astrophysical Journal*, 730, 137